ATTORNEY DOCKET NO. 2003028-0048 (Ariad 331D USD1)

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Appellant:	Schreiber & Crabtree	Examiner:	Vogel
Serial No.:	09/834,424	Art Unit:	1636
Filed:	April 13, 2001		
For:	METHODS AND MATERIALS INVOLVING DIMERIZATION-MEDIATED REGULATION OF BIOLOGICAL EVENTS		

Mail Stop Appeal Brief - Patents
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

Certificate of Mailing

I certify that this correspondence is being deposited with the United States Postal Service with sufficient postage as First Class Mail in an envelope addressed to Mail Stop Appeal Brief – Patents, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450

August 21, 2006
Date

Signature

Mary B. Wilson
Typed or Printed Name of person signing certificate

APPEAL BRIEF UNDER 37 C.F.R. § 1.192

Appellant appeals to the Board of Patent Appeals and Interferences (the "Board") from the Examiner's rejection of claims 8-29. A notice to this effect was filed pursuant to 37 C.F.R. § 1.191(a) on January 13, 2006 and received by the Patent Office on January 23, 2006.

Filed herewith is a Petition under 37 C.F.R. § 1.136 for a five (5) month extension of time, from March 23, 2006, up to and including August 23, 2006, to file this Appeal Brief (the "Brief"). This Appeal Brief is therefore timely filed on August 21, 2006. Pursuant to 37 C.F.R. § 1.192(a), this Brief is being filed in triplicate.

Also enclosed is a credit card payment form to cover the $2160.00 fee under 37 C.F.R. § 1.17 (a)(5) for the Petition and the $500.00 fee under 37 C.F.R. § 41.20(b)(2) for the Appeal Brief. Please charge any additional fees (or credit any overpayment), to our Deposit Account 03-1721.

Real Parties in Interest

As a result of assignments by the inventors in parent application U.S. Serial No. 09/430,508 (filed October 29, 1999), the real parties in interest in this application are the President and Fellows of Harvard College ("Harvard") and the Board of Trustees of Leland S. Stanford Junior University ("Stanford"). The assignments to Harvard and Stanford were recorded in the Patent and Trademark Office on March 20, 2000 at Reel 010691, Frames 0402

and 0291, respectively.

Related Appeals and Interferences

Appellant has filed an Appeal Brief for co-pending application U.S. Serial No. 09/430,508 that addresses some issues that overlap with the issues presented here. No other appeals or interferences are known to the Appellant, the Appellant's legal representative, or the Appellant's assignee that will directly affect or be directly affected by the Board's decision in this appeal. Similarly, no such appeals or interferences are known that may have a bearing on the Board's decision in this appeal.

Status of Claims

The application was filed with claims 1-7. Claims 1-7 were canceled in an Amendment filed April 1, 2003; claims 8-29 were added. Claims 8 and 18 were amended in an Amendment filed April 22, 2005. Claims 8-29 were finally rejected in an Office Action mailed July 15, 2005. The rejection of claims 8-29 is hereby appealed. A listing of pending claims 8-29 is provided as **Attachment I**.

Status of Amendments

There are no outstanding amendments to the claims.

Summary of Invention

Dimerization (or generally oligomerization) of proteins is a biological control mechanism that contributes to the activation of cell surface receptors, transcription factors, vesicle fusion proteins and other classes of intra- and extracellular proteins. The present invention is directed to methods that regulate the dimerization of such endogenous proteins using non-peptidic "dimerizing" agents. Pending claims 8-29 relate to *methods for preparing* these "dimerizers". One such method involves preparing a dimerizer which includes a first non-peptidic moiety that binds to one of the protein mediators covalently linked with a second non-peptidic moiety that binds to the other protein mediator. The resulting dimerizer binds to both protein mediators. Claim 8 is drawn to embodiments in which the protein mediators are cell-surface receptors. Claim 19 is drawn to embodiments in which the protein mediators are different and the first and second moieties of the dimerizer are different. Claims 9-18 and 20-29 depend from claim 8 and/or claim 19.

Issues

The issues on appeal are:

(1) Are claims 8-29 invalid for lack of written description?

(2) Are claims 8-27 anticipated by Wold?

(3) Are claims 8-29 anticipated by Ji?

Grouping of Claims

The claims stand or fall together for issue numbered (1) above, as indicated below:

(1) Claims 8, 9, 12, 19, 20, 21, 28 and 29 stand or fall together.

(2) Claim 10 stands or falls alone.

(3) Claim 11 stands or falls alone.

(4) Claim 13 stands or falls alone.

(5) Claim 14 stands or falls alone.

(6) Claim 15 stands or falls alone.

(7) Claim 16 stands or falls alone.

(8) Claim 17 stands or falls alone.

(9) Claim 18 stands or falls alone.

(10) Claims 22-24 stand or fall together.

(11) Claims 25-26 stand or fall together.

(12) Claim 27 stands or falls alone.

The claims stand or fall together for issue numbered (2) above, as indicated below:

(1) Claims 8-16 and 18-27 stand or fall together.

(2) Claim 17 stands or falls alone.

The claims stand or fall together for issue numbered (3) above, as indicated below:

(1) Claims 8-16 and 18-27 stand or fall together.

(2) Claim 17 stands or falls alone.

(3) Claims 28-29 stand or fall together.

Argument

The Pending Claims are not Invalid for Lack of Written Description

The pending claims stand rejected for lack of written description. In particular, the Examiner states that the disclosure in the specification does not reasonably convey to one skilled in the art that the inventors had possession of the claimed invention at the time the application was filed. This rejection is respectfully traversed; reconsideration and withdrawal is requested.

The written description requirement imposes a duty on patent applicants to notify the public of the scope and content of their inventions. The requirement is satisfied if one skilled in the art would reasonably conclude that the inventors were in possession of the claimed invention at the time the patent application was filed. *Vas-Cath, Inc. v. Mahurkar*, 935 F.2d 1555 (Fed. Cir. 1991). See also Guidelines for Examination of Patent Applications under the 35 U.S.C. § 112, ¶ 1, "Written Description" Requirement, 66 Fed. Reg. 4, 1099 (2001). A determination of whether the written description requirement is satisfied requires reading the disclosure in light of the knowledge possessed by those skilled in the art at the time that the application was filed. *In re Alton*, 76 F.3d 1168 (Fed. Cir. 1996).

This application claims methods for preparing certain non-peptidic "dimerizing" agents that effect biological events mediated by the association of protein mediators. The "dimerizing" agents bind and thereby cause the protein mediators to associate.

As a preliminary matter, Appellant has previously noted that, as set forth in the disclosure (see for example page 9, lines 6-8), it was well established in the art at the time the application was filed that the relevant biological events *necessarily occur* whenever the appropriate proteins are associated. Only imprecise association is required. Appellant has therefore argued that a description of *binding* is by necessity also a description of *effecting*.

To support this argument, Appellant previously submitted a reference by Spaargaren et al. demonstrating (in 1991, see Exhibit A) that a variety of different bivalent antibodies to the EGF receptor all successfully activated the receptor. The different antibodies recognized different epitopes on the receptor and therefore presumably bound in different ways from one another, as well as from its natural ligand. However, any binding that accomplished oligomerization also achieved activation. Previously cited references by Watowich et al. (see page 9, lines 8-12, Exhibit B) and Fuh et al. (see page 9, lines 12-15, Exhibit C) similarly

demonstrated (in 1992) activation of EPO receptors and GH receptors, respectively, through association achieved by engineered disulfide bonds or bivalent antibodies. In each of these references, association, *regardless of how achieved*, effected a biological event. As further evidence that this correlation was recognized in the art, Appellant also provided a 1994 article by Austin et al. (see Exhibit D) that states (in the context of the present invention):

> "It is easy to imagine the mixing and matching of different protein-binding surfaces using synthetic organic chemistry, to create new dimerizers with tailor made properties. Since protein dimerizers simply create a high local concentration of a particular protein at a particular cellular location, *their actions will not require the geometric precision associated with allosteric agents*" (see page 136, *emphasis* added).

Thus, Appellant respectfully submits that association of protein mediators triggers their activation and therefore, for the purposes of the present written description question, a description of *binding* is necessarily a description of *effecting* in the context of the present claims.

In the most recent Office Action mailed July 15, 2005, the Examiner disputes this argument (see paragraph spanning pages 4-5). Specifically, the Examiner argues that the references only "disclose [...] activation for particular instances" and that this is insufficient to show that "all binding 'is necessarily a description of *effecting*'." Appellant respectfully submits that the references provide a far stronger correlation between binding and effecting than the Examiner suggests. Indeed, the references demonstrate that a range of different receptors (e.g., EGF receptors, EPO receptors, GH receptors, etc.) can be successfully activated by simple association and that this activation is independent of the nature of the association (e.g., it works with bivalent antibodies, covalent cross-linking via disulfide bridges, etc.). Appellant respectfully submits that this strong correlation between binding and effecting must have a significant impact on the amount and nature of the description necessary to satisfy the written description requirement.

The Examiner also challenges Appellant's reliance on the aforementioned quote from Austin arguing that Austin does not state that "all and any agents that bind to a receptor, necessarily trigger a biological effect" (see page 4 of Office Action mailed July 15, 2005). The Examiner then implies that Austin did not recognize the correlation but merely "speculates that

one can create dimerizers with 'tailor made properties' that 'simply create a high local concentration of a particular protein at a particular cellular location'" (see top of page 5 of Office Action mailed July 15, 2005). Appellant respectfully submits that when viewed in their full context, the quoted statements in Austin are far more powerful than the Examiner suggests. Indeed, the whole premise of the Austin reference is that by targeting proteins that are activated by "dimerization," the inventive "dimerizing" agents need *only* create a high local concentration of a particular protein at a particular cellular location *in order to trigger a biological event* (e.g., see page 131 of Exhibit D). As explained in the clause that directly follows the section quoted by the Examiner, this is in stark contrast with the more stringent requirements of "allosteric" agents that trigger biological events by inducing allosteric (i.e., conformational) changes in the structures of a protein mediator ("[s]ince protein dimerizers simply create a high local concentration of a particular protein at a particular cellular location, *their actions will not require the geometric precision associated with allosteric agents*"). Appellant respectfully submits that Austin clearly supports the position that, for the purposes of the present invention, a description of *binding* is by necessity also a description of *effecting*.

Given that *binding* and *effecting* are simultaneously described in the present case, Appellant has urged that the written description standard articulated explicitly for antibodies is equally applicable here. That is, in *Noelle v. Lederman*, 355 F.3d 1343 (Fed. Cir. 2004), the Federal Circuit held that "as long as an applicant has disclosed a 'fully characterized antigen,' either by its structure, formula, chemical name, or physical properties, or by depositing the protein in a public depository, *the applicant can claim an antibody by its binding affinity to that described antigen*" (*emphasis* added). *Id.* at 1349. Thus, when the antigen is fully defined, the antibody is also described. Echoing the USPTO Written Description guidelines, the court listed three factors to support this conclusion, namely "the well defined structural characteristics for the five classes of antibody, the functional characteristics of antibody binding, and the fact that the antibody technology is well developed and mature." *Id.* In *Noelle*, the patent applicant ultimately failed to satisfy the written description requirement because he had failed to isolate and thereby characterize the antigen. *Id.*

In the present case, there is no dispute that the target is fully defined (e.g., see the description of representative protein mediators on pages 7-11 and references cited therein). Appellant further submits that consideration of the three *Noelle* factors reveals the strong

parallels between the present situation and the antibody case addressed in *Noelle*.

With respect to the first factor, it is correct that antibodies, as a class of molecules, have well defined structural characteristics and structural similarities. However, these characteristics and similarities are irrelevant to their binding capabilities. In fact, the portion of an antibody that is responsible for its binding attributes is referred to as the *variable* portion precisely because its amino acid sequence *differs* from that of other antibodies. It is not possible in advance to determine which variable region sequences will allow an antibody to bind to a particular antigen. The relevant structural characteristics of antibodies are therefore no more well defined than those of the claimed oligomerizing agents. In this context, the Examiner stated in the most recent Office Action that the written description requirement is not satisfied in this case because "the skilled artisan cannot envision the chemical structure of the encompassed genus of agents which will bind to two cell surface receptors in a manner which results in an effect on a biological event mediated by the association of said receptors, and therefore conception is not achieved until reduction to practice has occurred, regardless of the complexity or simplicity of the method of isolation or identification" (see bottom of page 5 of Office Action mailed July 15, 2005). Appellant respectfully submits that this argument cannot be reconciled with the holding in *Noelle*. Indeed, for the aforementioned reasons, a skilled artisan could not have "envisioned the chemical structure of a genus of antibodies" without making them and yet the court in *Noelle* would have been willing to allow claims to that very genus. It is therefore apparent that reduction to practice and precise chemical structures are not necessarily required in order to satisfy the written description requirement.

With respect to the second factor, the relevant "functional characteristics" in *Noelle* were the ability of the antibodies to bind to a known antigen. In this case, the "dimerizing" agents are non-peptidic agents that bind to a known target. The parallel is clear and inescapable.

With respect to the third factor, the *Noelle* court's reference to antibody technology as "well developed and mature" must rest on the conclusion that those of ordinary skill in the art could be confident in advance that, given a well characterized antigen, they would be able to identify and produce antibodies that bind to it. The same holds true in the present case. Technologies for identifying non-peptidic agents that bind to a given target were well established at the time the present application was filed. A variety of such agents were already known and available (see, for example, page 14 of the specification). As discussed in the 1994 review

article by Gordon et al. (see Exhibit E, especially pages 1390-1392), combinatorial libraries of agents were being developed (see also Bunin et al., Exhibit F and DeWitt et al., Exhibit G that are cited by Gordon et al.). A whole host of binding assays for screening these diverse agents were also known (see, for example, pages 15-19 of the specification). In fact, high-throughput screening systems for identification of binding agents were being used at the time the present application was filed. Appellant has previously submitted a review by Burch et al. (see Exhibit H) that discusses the state of high-throughput screening methods in 1991. The 1994 review article by Gordon et al. is also relevant (see Exhibit E, especially pages 1393-1397). Thus, at the time the present application was filed, the starting materials and technology for identifying oligomerizing agents that bind to a known, defined target, were both established. Furthermore, Appellant respectfully submits that a person skilled in the art would recognize in advance that for any given protein mediator many suitable oligomerizing agents will exist. Thus, as with antibodies, the universe of suitable oligomerizing agents that one has to choose from is large and there are in effect *multiple* correct answers to the same question. The practitioner need only identify one of these answers. This is in stark contrast to inventions involving a single gene, protein or allosteric agent where the search is for a unique entity with a unique structure. In those cases there is only a *single* correct answer.

The Examiner has so far resisted application of the *Noelle* standard in this case, stating that "while methods of isolation may be known in the art, the existence of such methods of potentially isolating agents which meet the functional limitations of the claims is not sufficient to show that applicants had possession of the invention, i.e., invented what is now being claimed" (see page 6 of Office Action mailed July 15, 2005). Appellant respectfully submits that the Examiner has failed to address the fact that in *Noelle* the court explicitly held that possession of the invention can be based on the existence of such methods.

The Examiner may well be resisting application of the *Noelle* standard on the ground that the technology for making oligomerizing agents was arguably not as mature as the technology for making antibody binding agents. If so, Appellant would respectfully submit that the proper question is not whether the present specification describes a process that will necessarily make every oligomerizing agent encompassed by the claims, but rather whether a person of ordinary skill in the art, reading the specification, would understand that the inventors were "in possession of" the claimed invention. That is, would a person of ordinary skill in the art have understood

that the claimed invention could successfully be practiced, based on the specification? To this question, the answer is clearly "yes."

Indeed, Appellant notes that the specification includes references to a variety of binding agents (e.g., benzodiazepines, prostaglandins, beta-turn mimetics, alpha- and beta-blockers, etc. on page 14, lines 7-11 of the specification) that were available at the time the application was filed, and that were known to bind protein mediators of biological events. Collections of synthetic compounds and combinatorial libraries of compounds were also available (see page 19, lines 1-2 and Exhibits E-G). The specification also defines the characteristics and methods that could be used to test these and other agents for desirable binding ability (see, for example, pages 15-19). Thus, a huge number of useful agents were already known and available in the art; others could readily be identified as they came available. An oligomerizing agent for a protein mediator of interest could therefore be prepared by selecting a known binding agent or by screening available agents for binding against the protein mediator. No further guidance is required to describe possession of the invention.

In this context, the Examiner in parent application U.S. Serial No. 09/430,508 has characterized the process for preparing oligomerizing agents as a method of "identifying" as contrasted with the process for preparing antibodies which he characterized as a method of "making" ("identifying something is not the same as making something"). Should the Examiner in this case make a similar argument, Appellant would respectfully submit that this characterization is misleading. Indeed, while the process for preparing an oligomerizing agent might in some situations require a step of screening agents for binding to the protein mediator of interest, this is no different than the process for preparing antibodies. Indeed, the final step in any antibody preparation process involves screening hybridomas to identify those that produce antibodies against the antigen of interest. Thus, the mere fact that a screening or "identification" step might be involved in the present invention does not distinguish it from the situation in *Noelle*.

Finally, Appellant has previously provided objective *evidence* that those of ordinary skill in the art, presented with teachings of the type found in the present specification, immediately recognize their generality and breadth, even when only limited exemplification is provided. Specifically, Appellant has cited references by Qureshi and Tian, published after the filing of the present case, that report examples of oligomerizing agents of the type whose use is encompassed

by the present claims. Qureshi provides a single example of an oligomerizing agent that dimerized the EPO receptor and concluded that "most cytokine receptors can be ligated together in an active conformation by a nonpeptidyl molecule" (see Exhibit I, last sentence, page 12161); Tian tested a single oligomerizing agent that dimerized the G-CSF receptor and concluded that "a small molecule can trigger the activation of a large (~120 kD) receptor protein that requires dimerization for activation" (see Exhibit J, last sentence, page 259). In each case, a single example justified broad conclusions because those of ordinary skill in the art immediately understood it would work, that Qureshi and Tian were "in possession of" broad discoveries.

Of further note, both Qureshi and Tian identified their binding agents by screening libraries of known compounds using a high-throughput assay for each receptor (see page 12158 of Qureshi and page 257 of Tian). Qureshi screened known antagonists of the EPO receptor and then linked eight copies of a preferred antagonist together to form an oligomerizing agent. Tian screened libraries of small molecules directly for oligomerizing agents. While both agents were identified after the present application was filed, the routine methods used reinforce the fact that the identification of binding agents is no more burdensome than the preparation of antibodies. This was recognized in a 1999 article by Clemons commenting that Tian's work suggested "the synthetic attainability of a wide range of receptor dimerizers" and conclusively showed that synthetic nonpeptidyl ligands can mimic the effects of polypeptide growth factors (see Exhibit K, especially page 114).

The Examiner has repeatedly rejected Appellant's arguments with regard to Qureshi and Tian on the ground that the particular agents used by Qureshi and Tian were not described in the specification and therefore cannot be relied upon as evidence that the present specification teaches a sufficient number of representative examples. Appellant does not rely on Qureshi and Tian as evidence that the present specification teaches a sufficient number of representative examples. Appellant refers to the specification and knowledge in the prior art to make that point. Rather, Appellant relies on Qureshi and Tian to provide *objective* evidence that only a small number of specific examples (only one!) is required to describe the present invention to those of ordinary skill in the art. The present specification contains abundant description, and fully satisfies the written description requirement.

Independent claims 8 and 19 and dependent claims 9, 12, 20, 21 and 29 stand or fall together based on the aforementioned arguments. The following claims stand and fall separately

from these and other claims for the following reasons.

Claim 10 specifies that biological event is mediated by the association of cell-surface receptors for a growth factor, cytokine, or hormone. The universe of receptors whose oligomerization is achieved in the methods of claim 10, is smaller than that of claim 8, so that the level of description in the specification required would be reduced as compared with claim 8. Thus, even if claim 8 were not fully supported by the specification, claim 10 would be.

Claim 11 is even more specific than claim 10, naming particular receptors. Even if claims 8 and 10 were not supported, claim 11 would be.

Claim 18 lists a different set of receptors and is also separately patentable.

Claim 13 specifies that the oligomerizing agent binds to a cytoplasmic portion of the receptor; claim 14 specifies that it binds to an extracellular portion. The focus on a binding portion may simplify the identification of oligomerizing agents whose use falls within the scope of the claims. Thus, even if claim 1 were found not enabled, claims 13-14 would be. Furthermore, agents that bind to extracellular portions of a receptor could activate signal transduction without crossing the cell membrane, whereas agents that bind to cytoplasmic portions need to be cell permeant. Thus, claims 13 and 14 have separate bases for patentability.

Claim 15 recites methods in which the oligomerizing agent has a particularly high affinity for the receptor. The specification provides appropriate guidance for identifying and using that set of oligomerizing agents; claim 15 is separately patentable from the other claims.

Claim 16 recites methods in which the oligomerizing agent includes non-peptidic moieties that have a molecular weight less than 5 kD. The universe of non-peptidic moieties that have a molecular weight less than 5 kD is smaller than that encompassed by claim 8, so that the level of description required would be reduced as compared with claim 8. Thus, even if claim 8 were not fully supported by the specification, claim 16 would be.

Claim 17 recites methods in which the oligomerizing agent is membrane permeant. The specification provides appropriate guidance for identifying and using that set of oligomerizing agents; claim 17 is separately patentable from the other claims.

Claims 22-24 are limited to methods in which the biological event is transcriptional regulation driven by the association of a protein containing a DNA-binding domain and a protein containing a transcriptional regulatory domain. The universe of proteins whose oligomerization is achieved in the methods of claims 22-24, is different than that of claim 8 and smaller than that

of claim 19, so that the level of description in the specification required would be different or reduced as compared with claims 8 and 19, respectively. Thus, even if claims 8 and 19 were not fully supported by the specification, claims 22-24 would be.

Claims 25-26 are limited to methods in which the biological event is translocation of a selected protein to a predetermined cellular component driven by the association of the selected protein and a constituent of the predetermined cellular component. The universe of proteins whose oligomerization is achieved in the methods of claims 25-26, is different than that of claim 8 and smaller than that of claim 19, so that the level of description in the specification required would be different or reduced as compared with claims 8 and 19, respectively. Thus, even if claims 8 and 19 were not fully supported by the specification, claims 25-26 would be.

Claim 27 is limited to methods in which the biological event is destruction of a selected protein driven by the association of the selected protein and a constituent of the proteasome. The universe of proteins whose oligomerization is achieved in the method of claim 27, is different than that of claim 8 and smaller than that of claim 19, so that the level of description in the specification required would be different or reduced as compared with claims 8 and 19, respectively. Thus, even if claims 8 and 19 were not fully supported by the specification, claim 27 would be.

Claims 8-27 are not Anticipated by Wold

Claims 8-27 were rejected under 35 U.S.C. § 102(b) as being anticipated by Wold, Methods Enzymology 11:617-640, 1966 ("Wold"). Applicant respectfully submits that Wold does not anticipate the pending claims because it fails to teach each and every element of the claimed methods. MPEP § 2131.

Claims 8-18 recite methods that involve an agent that *binds* to two or more endogenous cell surface receptor molecules. Claims 19-27 recite methods that involve an agent that *binds* to two or more endogenous protein mediators. Appellant has previously argued that a skilled person in the art readily understands that the claim term "binds" is used in the art to refer to *non-covalent* associations (e.g., between an antibody and antigen or a receptor and a ligand). Appellant also noted that the explicit teachings of the application (e.g., see discussion of binding affinities on page 11, lines 13-18; discussion of exemplary receptor binding moieties on page 13, lines 19-25; discussion of affinity assays for identifying receptor binding moieties on pages 14-

19; etc.) conform with this interpretation and reinforce that, as used in the context of this application, the term "binds" refers to non-covalent associations.

Wold does not teach methods that involve an agent that *binds* non-covalently to two or more endogenous protein mediators. Instead, Wold teaches bifunctional reagents that *react with* and thereby form *covalent* bridges within or between proteins. As explained by Wold (page 617), these *covalent* bridges were used in the 1960s-1980s to study and determine the three dimensional structures of proteins and protein complexes. Exemplary reagents that were reviewed by Wold in 1966 include *N*-substituted maleimide derivatives that *react* with sulfydryl groups (pages 622-623); alkyl halides that *react* with sulfydryl groups, sulfides, imidazole, and amino groups (pages 623-627); aryl halides that *react* with amino, tyrosine phenolic, sulfydryl and imidazole groups (pages 627-632); etc. All of these reagents react with groups that are present within proteins to form covalent bonds. Based on the foregoing, Appellant has argued that the pending claims cannot be anticipated by Wold since it fails to teach an agent that *binds* to proteins.

In the most recent Office Action, the Examiner maintained this prior art rejection by arguing that "there is no [...] limitation in the claims to exclude covalent binding [...]; contrary to applicant's arguments, the term "binds" does not exclude covalent binding" (see page 3 of Office Action mailed July 15, 2005). Appellant respectfully disagrees with this line of reasoning. During patent examination, the pending claims must be "given the broadest *reasonable* interpretation *consistent with the specification*" (*emphasis* added). MPEP § 2111 citing from *In re Prater*, 415 F.2d 1393, 1404-05 (CCPA 1969). The "broadest reasonable interpretation" must also be consistent with the interpretation that those skilled in the art would reach. MPEP § 2111 citing *In re Cortright*, 165 F.3d 1353, 1359 (Fed. Cir. 1999).

While this analysis must avoid importing limitations from the specification into the claims; it is entirely appropriate to interpret limitations that are already explicitly recited in the claim based on the specification. Specifically, in *Prater*, the court explained that "reading a claim in light of the specification, to thereby interpret limitations explicitly recited in the claim, is a quite different thing from 'reading limitations of the specification into a claim,' to thereby narrow the scope of the claim by implicitly adding disclosed limitations which have no express basis in the claim". *Id.* Here, we are asking the Examiner to do the former permissible analysis, namely to interpret the explicitly recited claim term "binds" in light of the specification. As

previously argued, the specification makes it clear to the skilled person that the term "binds" refers to non-covalent association. For example, on page 11, lines 13-18, the specification introduces the "dimerizing" agents by describing their binding properties in terms of binding *affinities* (K_d below about 10^{-6}, more preferably below about 10^{-7}, 10^{-8} or 10^{-9}). A skilled person would appreciate that binding affinities are only used in reference to non-covalent associations, not covalent bonds. The discussion of *affinity* assays for identifying receptor binding moieties on pages 14-19 and the discussion of exemplary receptor binding moieties on page 13, lines 19-25 (all non-peptidic moieties that are known to bind non-covalently to proteins) further reinforces that the claim term "binds" refers to non-covalent associations, not covalent bonds. Finally, Appellant notes that the specification and claims explicitly use the terms "covalently linked", "covalently joined" and "covalently attached" to describe the covalent bond between the first and second non-peptidic moieties of the agent (e.g., see page 4, lines 25-28; page 13, lines 7-9 and 15-17; and claims 8 and 19). A skilled person would appreciate that this further differentiates the claim term "binds" from covalent bonding. For all of these reasons, Appellant respectfully requests withdrawal of this rejection.

For purposes of this rejection, claims 8-16 and 18-27 stand and fall together based on the aforementioned arguments. The following claim stands and falls separately from these claims for the following additional reasons. Claim 17 specifies that the "dimerizing" agent is membrane permeant. The Examiner has failed to point to any teaching in Wold that satisfies this limitation.

Claims 8-29 are not Anticipated by Ji

Claims 8-29 were rejected under 35 U.S.C. § 102(b) as being anticipated by Ji, Methods Enzymology 91:580-609, 1983 ("Ji"). Applicant respectfully submits that Ji does not anticipate the pending claims because it fails to teach each and every element of the claimed methods. MPEP § 2131. The reasoning here is identical to that presented above with respect to Wold.

As was the case with Wold, Ji does not teach methods that involve an agent that *binds* to two or more endogenous protein mediators. Instead, Ji teaches bifunctional reagents that *react with* and thereby form *covalent* bridges within or between proteins. As explained by Ji (page 580), these *covalent* bridges were used in the 1960s-1980s to study and determine the three dimensional structures of proteins and protein complexes. Ji specifically describes covalent crosslinking reagents that were developed after Wold's review. Thus, Ji describes covalent

reagents that react with specific protein groups (page 591-601) and covalent reagents that include photoactivable (pages 602-605) or cleavable groups (page 606-607). Nowhere does Ji remedy the deficiencies of Wold by teaching methods for preparing an agent that *binds* to two or more endogenous protein mediators. For these reasons and those discussed above with respect to Wold, Appellant respectfully requests withdrawal of this rejection.

For purposes of this rejection, claims 8-16 and 18-27 stand and fall together based on the aforementioned arguments. The following claims stand and fall separately from these claims for the following additional reasons.

Claim 17 specifies that the "dimerizing" agent is membrane permeant. The Examiner has failed to point to any teaching in Ji that satisfies this limitation.

Claims 28 and 29 specify that the "dimerizing" agent is further mixed with a pharmaceutically acceptable carrier and optionally with one or more pharmaceutically acceptable excipients. The Examiner has failed to point to any teaching in Ji that satisfies this limitation.

Conclusion

Appellant again concludes with the belief that claims 8-29 are not anticipated by the prior art and are well supported by the specification. Allowance of the pending claims is earnestly requested.

Respectfully submitted,

Dated: August 21, 2006

Charles E. Lyon, D.Phil.
Reg. No. 56,630

PATENT DEPARTMENT
CHOATE, HALL & STEWART LLP
Two International Place
Boston, Massachusetts 02110
Telephone: (617) 248-5000
Facsimile: (617) 248-4000

4111119_1.DOC

Attachment I

to

Appeal Brief under 37 C.F.R. § 1.192

Pending Claims

Pending Claims

1-7. **(Canceled)**

8. **(Previously presented)** A method for preparing an agent that effects a biological event mediated by the association of two or more endogenous cell surface receptor molecules, the method comprising preparing an agent which includes a first non-peptidic moiety that binds to one of the cell surface receptor molecules covalently linked to a second non-peptidic moiety that binds to the other cell surface receptor molecule, wherein the agent binds to both cell surface receptor molecules.

9. **(Previously presented)** The method of claim 8 wherein the biological event is mediated by the association of two or more molecules of the same cell surface receptor and the first and second non-peptidic moieties are the same.

10. **(Previously presented)** The method of claim 9 wherein the cell surface receptor is a receptor for a cytokine, growth factor or other hormone.

11. **(Previously presented)** The method of claim 10 wherein the cell surface receptor is a receptor for erythropoietin ("EPO"), granulocyte colony stimulating factor ("G-CSF"), thrombopoietin ("TPO"), growth hormone ("GH"), interleukin-2 ("IL-2"), interferon-alpha, interferon-beta, or a neurotropic factor.

12. **(Previously presented)** The method of claim 8 wherein the biological event is mediated by the association of molecules of two different cell surface receptors and the first and second moieties are different.

13. **(Previously presented)** The method of claim 8 wherein the first and second non-peptidic moieties bind to cytoplasmic portions of the cell surface receptors.

14. **(Previously presented)** The method of claim 8 wherein the first and second non-peptidic moieties bind to extracellular portions of the cell surface receptors.

15. **(Previously presented)** The method of claim 8 wherein the agent binds to the cell surface receptors with a $K_d \leq 10^{-6}$ M.

16. **(Previously presented)** The method of claim 8 wherein the first and second non-peptidic moieties have a molecular weight less than 5 kD.

17. **(Previously presented)** The method of claim 8 wherein the agent is membrane permeant.

18. **(Previously presented)** The method of claim 8 wherein the cell surface receptors are selected from the group consisting of:
epidermal growth factor-receptor (EGF-R),
ataxia telangiectasia and rad-related 2/neuroblastoma oncogene (ATR2/neu),
hermaphrodite homolog-2/neuroblastoma oncogene (HER2/neu),
hermaphrodite-3/cellular-erythroblastic leukemia oncogene homolog-3 (HER3/c-erbB-3),
Xiphophorus melanoma receptor tyrosine kinase homolog (Xmrk);
insulin-like growth factor-I-receptor (IGF-1-R),
insulin receptor-related receptor (IRR);
platelet-derived growth factor receptor-a (PDGF-R-a),
platelet-derived growth factor receptor-ß (PDGF-R-ß),
colony stimulating factor-1-receptor (CSF-1-R, macrophage-colony stimulating factor-receptor (M-CSF-R)/cellular-McDonough feline sarcoma homolog (c-Fms)),
c-kit (Steel Factor Receptor, mast/stem cell growth factor receptor, HZ4-feline sarcoma virus viral oncogene homolog),
serine/threonine kinase/fms-like tyrosine kinase 2 (STK-1/Flk-2);
fibroblast growth factor-receptor (FGF-R),
[acidic-] fibroblast growth factor-receptor-1 (flg),
[basic-] fibroblast growth factor-receptor-2 (bek);

neurotrophic tyrosine kinases;
cell-surface determinant-3-z (CD3-zeta),
cell surface/class II determinant-3-h (CD3-eta);
ß chain of Fc/IgE receptor-1 (FCERI),
g chain of Fc/IgE receptor-1 (FCERI);
g chain of Fc receptor/cell-surface determinant-16 (Fcg-RIII/CD16);
cell-surface determinant-3-g (CD3-gamma) subunit,
cell-surface determinant-3-d (CD3-delta) subunit,
cell-surface determinant-3-e (CD3-epsilon) subunit;
Ig-a subunit of B-cell antigen receptor complex/membrane-bound,
Ig-associated protein-1 (Ig-a/MB-1),
Ig-ß subunit of B-cell antigen receptor complex/c membrane-bound,
Ig-associated glycoprotein B29 (Ig-ß/B29);
the common ß subunit shared by the granulocyte/macrophage-colony stimulating factor (GM-CSF), interleukin-3 (IL-3) and interleukin-5 (IL-5) receptors;
the ß chain of glycoprotein MW 130 KD (gp130) associated with the interleukin-6 (IL-6), leukemia inhibitory factor (LIF), ciliary neurotrophic factor (CNTF), oncostatin M, and interleukin-11 (IL-11) receptors;
the interleukin-2 (IL-2) receptor gamma subunit associated also with receptors for interleukin-4 (IL-4), interleukin-7 (IL-7) and interleukin-13 (IL-13);
the ß chain of the interleukin-2 (IL-2) receptor;
interferon (IFN) a receptor,
interferon (IFN) ß receptor,
interferon (IFN) g receptor;
growth hormone (GH) receptor,
erythropoietin (EPO) receptor,
prolactin receptor; and
the Transforming growth factor-ß (TGF-ß) family of cell surface receptors.

19. **(Previously presented)** A method for preparing an agent that effects a biological event mediated by the association of two or more endogenous protein mediators, the method

comprising preparing an agent which includes a first non-peptidic moiety that binds to one of the protein mediators covalently linked with a second non-peptidic moiety that binds to the other protein mediator, wherein the agent binds to both protein mediators, the biological event is mediated by the association of molecules of two different protein mediators and the first and second moieties are different.

20. **(Previously presented)** The method of claim 19 wherein at least one of the protein mediators is a cell surface receptor.

21. **(Previously presented)** The method of claim 19 wherein the two different protein mediators are cell surface receptors.

22. **(Previously presented)** The method of claim 19 wherein the biological event is transcriptional regulation, the first non-peptidic moiety binds to a protein containing a DNA-binding domain and the second non-peptidic moiety binds to a protein containing a transcriptional regulatory domain.

23. **(Previously presented)** The method of claim 22 wherein the transcriptional regulatory domain is a transcriptional activation domain.

24. **(Previously presented)** The method of claim 22 wherein the transcriptional regulatory domain is a transcriptional repression domain.

25. **(Previously presented)** The method of claim 19 wherein the biological event is translocation of a selected protein to a predetermined cellular component, the first non-peptidic moiety binds to the selected protein and the second non-peptidic moiety binds to a constituent of the predetermined cellular compartment.

26. **(Previously presented)** The method of claim 25 wherein the first non-peptidic moiety binds to a protein that functions only in the cytoplasm and the second non-peptidic moiety binds to a constituent of the nucleus or mitochondrion.

27. **(Previously presented)** The method of claim 19 wherein the biological event is destruction of a selected protein, the first non-peptidic moiety binds to the selected protein and the second non-peptidic moiety binds to a constituent of the proteasome.

28. **(Previously presented)** The method of claim 8 or 19 further comprising mixing the agent with a pharmaceutically acceptable carrier and optionally with one or more pharmaceutically acceptable excipients.

29. **(Previously presented)** A method which comprises providing an agent prepared according to the method of claim 8 or 19 and mixing the agent with a pharmaceutically acceptable carrier and optionally with one or more pharmaceutically acceptable excipients.

EXHIBIT A

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 266, No. 3, Issue of January 25, pp. 1733-1739, 1991
Printed in U.S.A.

Antibody-induced Dimerization Activates the Epidermal Growth Factor Receptor Tyrosine Kinase*

(Received for publication, August 6, 1990)

Marcel Spaargaren‡§¶, Libert H. K. Defize‡, Johannes Boonstra§, and Siegfried W. de Laat‡

From the ‡Hubrecht Laboratory, Netherlands Institute for Developmental Biology, Uppsalalaan 8, 3584 CT Utrecht, The Netherlands and the §Department of Molecular Cell Biology, University of Utrecht, Padualaan 8, 3584 CH Utrecht, The Netherlands

The relationship between epidermal growth factor receptor (EGF-R) protein tyrosine kinase activation and ligand-induced receptor dimerization was investigated using several bivalent anti-EGF-R antibodies directed against various receptor epitopes. In A431 membrane preparations and permeabilized cells, all antibodies were able to activate the EGF-R tyrosine kinase, as measured by EGF-R autophosphorylation and phosphorylation of other substrates on tyrosine residues. EGF-R tyrosine kinase activation correlated strongly with the induction of EGF-R dimerization. (i) Both processes specifically occurred in a narrow antibody concentration range; (ii) both processes required the presence of detergent; and (iii) both processes depended on antibody bivalence since monovalent Fab fragments were inactive yet regained full activity after cross-linking by a second bivalent antibody. These data demonstrate that antibody bivalence is essential and sufficient for EGF-R activation and that activation occurs regardless of the EGF-R epitope recognized. Finally, EGF-R dimerization was shown not to depend on receptor autophosphorylation since it still occurred in the absence of ATP. Also, partial inhibition of the tyrosine kinase activity by the specific EGF-R tyrosine kinase inhibitor tyrphostin AG 213 did not affect formation of EGF-R dimers. Taken together these results demonstrate that induction of EGF-R dimerization is sufficient and in case of antibody action, essential, for activation of the EGF-R tyrosine kinase and thus provide strong support for an intermolecular mechanism of EGF-R tyrosine kinase activation.

The epidermal growth factor receptor (EGF-R)[1] is a transmembrane glycoprotein of 170 kDa with an extracellular ligand binding domain, a single hydrophobic transmembrane stretch, and an intracellular domain containing protein tyrosine kinase activity (1). Binding of EGF to the receptor enhances this activity, thus catalyzing tyrosine phosphorylation of several protein substrates including the EGF-R itself (2). Activation of the EGF-R tyrosine kinase results in a cascade of biochemical and physiological responses, finally leading to stimulation of DNA synthesis and cell division in most cells (3). Using EGF-R mutants, the tyrosine kinase activity was shown to be essential for stimulation of all known early and late responses involved in mitogenic signaling (4–9).

The mechanism whereby EGF binding activates the EGF-R tyrosine kinase still remains to be elucidated. Two models have been proposed. According to the intramolecular activation model, EGF binding leads to a conformational change in the extracellular EGF-R domain which is transmitted through the hydrophobic transmembrane stretch to the cytoplasmic tyrosine kinase domain, resulting in activation of the EGF-R tyrosine kinase (10–13). This model implies that EGF-R monomers can act as transmembrane signal transducers. According to the intermolecular activation model, EGF causes a shift in a hypothetical equilibrium between inactive EGF-R monomers toward active dimers, resulting in enhanced tyrosine kinase activity (14, 15). However, using Triton X-100-solubilized purified EGF-R material, Das and colleagues (16–18) interpreted their data as EGF binding resulting in dissociation of inactive EGF-R oligomers, thus shifting the equilibrium toward catalytically active monomers. In accordance with an intermolecular, active dimer, activation model, a number of data show that EGF-induced dimerization occurs. Yarden and Schlessinger (19, 20) have demonstrated a reversible EGF-R oligomerization after EGF treatment of solubilized immunoaffinity-purified EGF receptors. EGF-induced dimerization of receptors could also be detected by chemical cross-linking of receptors in Triton X-100-solubilized receptor preparations (21, 22), membrane preparations (23), and intact cells (24–26). In several cases enhanced autophosphorylation in EGF-R dimer complexes has been reported (21, 22). Taken together, these data indicate that EGF binding results in EGF-R dimerization; however, a direct cause and effect relationship between EGF-R dimerization and activation of the tyrosine kinase has not been established yet.

To study the relationship between EGF-R dimerization and tyrosine kinase activation, we made use of anti-EGF-R antibodies. It has been reported that under certain conditions a number of such antibodies can mimic EGF in activating the EGF-R tyrosine kinase (19, 27–32). In view of the intermolecular activation model, it has been proposed that this ability is due to the bivalent nature of such antibodies (19, 30, 31). In the present study we examined the relationship among antibody bivalence, EGF-R dimerization, and EGF-R tyrosine kinase activity in both membrane preparations and permeabilized cells. We demonstrate that induction of receptor

* This work was supported by the Center for Developmental Biology, Utrecht, The Netherlands. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

¶ To whom correspondence should be sent. Tel.: 31-30-510211; Fax: 31-30-516464.

[1] The abbreviations used are: EGF-R, epidermal growth factor receptor; EGF, epidermal growth factor; PBS, phosphate-buffered saline; SDS, sodium dodecyl sulfate; PAGE, polyacrylamide gel electrophoresis; Hepes, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; EDAC, 1-ethyl-3-(3-(dimethylamino)propyl)carbodiimide; RAM, rabbit anti-mouse antiserum; SwAM, swine anti-mouse antiserum; RIPA, radioimmune precipitation buffer.

dimerization is sufficient and in case of antibody action, essential, for stimulation of EGF-R autophosphorylation and tyrosine kinase activation. This establishes a causative role for EGF-R dimerization in tyrosine kinase activation and thus supports an intermolecular activation model.

EXPERIMENTAL PROCEDURES

Materials—EGF (receptor grade) was obtained from Collaborative Research (Waltham, MA); [γ-^{32}P]ATP and Na^{125}I, from Amersham Corp.; protein A-Sepharose, from Pharmacia LKB Biotechnology Inc.; plastic foil thin layer cellulose plates (5577), from Merck; and rabbit anti-mouse as well as swine anti-mouse IgG/7 S antiserum were from Nordic (Tilburg, The Netherlands). Monoclonal antibody 2E9, 2D11, and polyclonal antiserum 281–7 were prepared and IgG purified as described previously (26, 29). Antibodies 528 and 225 (31) were a kind gift of Dr. J. Mendelsohn (Sloan-Kettering Institute for Cancer Research, New York); for later experiments, 528 was obtained from Oncogene Science.

Cell Culture and Preparation of Membranes—A431 cells were grown in Dulbecco's modified Eagle's medium supplemented with 7.5% fetal calf serum and buffered with 40 mM $NaHCO_3$ under a 7.5% CO_2 atmosphere. Membrane preparations were derived from A431 cells by vesiculation essentially as described (33).

Preparation of Monovalent Fab Fragments—Fab fragments were prepared by papain digestion. Papain (10%, w/v) was preactivated in 0.1 M phosphate buffer (pH 7.6), containing 2 mM EDTA and 10 mM cysteine at 37 °C for 30 min. 2E9 (1 mg/ml) was digested with 1% papain in 0.1 M phosphate buffer (pH 7.6) in the presence of 2 mM EDTA and 10 mM cysteine for 16 h at 37 °C. The reaction was stopped by the addition of 10 mM iodoacetamide at 37 °C for 20 min. The digest was dialyzed against Ca^{2+}- and Mg^{2+}-free PBS. For purification, the preparation was applied on a 5-ml protein A-Sepharose column (Pharmacia), and Fab fragments (unbound fraction) were collected and concentrated in a collodion bag (Sartorius) under reduced external pressure. The digestion and quality of the Fab fragments were tested by SDS-PAGE and by immunoprecipitation of EGF-R in the presence or absence of a second antibody (RAM).

EGF-R Immunoprecipitation—After surface iodination or phosphorylation reactions, cells were washed with PBS free of Ca^{2+} and Mg^{2+} and lysed for 15 min on ice in 500 μl of RIPA containing 150 mM NaCl, 20 mM Tris/HCl (pH 8.0), 10 mM NaH_2PO_4, 5 mM EDTA, 100 μM Na_3VO_4, 10% glycerol, 1% Nonidet P-40, 1% sodium deoxycholate, 0.1% SDS, 1 mM phenylmethylsulfonyl fluoride, and 100 kallikrein inhibitor units/ml aprotinin (membrane preparations were directly lysed in RIPA). The lysates were clarified by centrifugation at 15,000 × *g* for 15 min in an Eppendorf centrifuge. Anti EGF-R antibody 528 (2 μg) or antiserum 281–7 (10 μl) bound to protein A-Sepharose was added to the lysates, and the mixtures were rotated 1 h at 4 °C. Precipitates were washed once with RIPA, twice with 20 mM Tris/HCl (pH 7.4), 0.5 M NaCl, 10% glycerol, and 1.5% Nonidet P-40, and twice with the same buffer containing 0.15 M NaCl. The precipitates were boiled in Laemmli sample buffer and applied to gel.

EGF-R Autophosphorylation and Exogenous Substrate Phosphorylation Measurements in Membrane Preparations—Analysis of phosphorylation of the EGF-R and the tyrosine-containing peptide angiotensin I was performed as described (34) except that the phosphorylation reactions contained 5 μCi of [γ-^{32}P]ATP and were allowed to proceed for 10 min.

EGF-R Autophosphorylation and Endogenous Substrate Phosphorylation Measurements in Saponin-permeabilized Cells—Cell permeabilization on A431 cells grown to subconfluence in 12-well tissue culture clusters (Costar) was performed essentially as described by Guigni *et al.* (35). Briefly, after preincubation in Dulbecco's modified Eagle's medium/Hepes (10 mM, pH 7.5) with or without 100 μg/ml Fab fragments for 2 h on ice as indicated, followed by two washes with Dulbecco's modified Eagle's medium/Hepes, A431 cells were incubated with 2E9 (50 μg/ml) or RAM (500 μg/ml) for 20 min at room temperature in Dulbecco's modified Eagle's medium/Hepes. After this incubation, cells were washed two times with cold PBS and once with cold permeabilization buffer containing 20 mM Hepes (pH 7.4), 145 mM NaCl, 5.4 mM KCl, 1 mM $CaCl_2$, 2 mM $MgCl_2$, 2 mM $MnCl_2$, 20 μM Na_3VO_4, 10 mM *p*-nitrophenyl phosphate, 1 mM phenylmethylsulfonyl fluoride. Phosphorylation was started by the addition of 250 μl of permeabilization buffer containing 0.005% saponin and 15 μM [γ-^{32}P]ATP (5 μCi) and EGF as indicated. The phosphorylation reaction was allowed to proceed for 10 min on ice and was stopped by washing twice with cold Ca^{2+}- and Mg^{2+}-free PBS containing 200 μM Na_3VO_4 followed by the addition of sample buffer (endogenous substrate phosphorylation) or by lysing the cells in RIPA in order to immunoprecipitate the EGF-R (autophosphorylation). After immunoprecipitation of the EGF-R, receptor phosphorylation was analyzed by 5–15% SDS-PAGE, autoradiography, and excision of EGF-R containing gel bands for measurement of radioactivity. Substrate phosphorylation was measured after 5–15% SDS-PAGE by scanning of the autoradiograph on a Joyce-Loebl Chromoscan 3 apparatus.

Two-dimensional Phosphoamino Acid Analysis—Two-dimensional phosphoamino acid analysis of immunoprecipitated EGF-R was performed as described (26).

EGF-R Dimer Cross-linking in Membrane Preparations—A431 membranes (20 μg) were incubated at room temperature for 20 min with EGF (1 μg/ml), 2E9 (50 μg/ml), or RAM (500 μg/ml) after a 2-h preincubation on ice with 100 μg/ml Fab fragments as indicated. [γ-^{32}P]ATP-containing phosphorylation mix was added for subsequent incubation for 10 min on ice, and phosphorylation was stopped by the addition of 10 mM EDTA. After the phosphorylation reaction, 15 mM cross-linker EDAC was added, followed by incubation for 15 min at 30 °C, sample buffer was added, and samples were heated for 5 min at 90 °C to stop the reaction. Samples were applied to 3–10% SDS-PAGE, and receptor dimerization was quantitated by liquid scintillation measurements of the excised monomer, dimer, and background gel bands.

Surface Iodination and EGF-R Dimer Cross-linking in Intact Cells—A431 cells were grown to subconfluence in 12-well tissue culture clusters (Costar) in the presence or absence of 100 μM tyrphostin AG 213 for the last 16 h as indicated, washed extensively with PBS, and surface iodinated essentially as described (25). Briefly, cells were incubated with 300 μl of PBS containing 6 μg of lactoperoxidase (Boehringer Mannheim), 0.06 units of glucose oxidase, 100 μCi of Na^{125}I, and iodination was initiated by the addition of 20 mM glucose. After 30 min of surface iodination at room temperature on a gently shaking platform, cells were washed three times with PBS and incubated with EGF (500 ng/ml) for 1 h on ice in the presence or absence of 100 μM tyrphostin AG 213 as indicated, washed twice with cold PBS, and incubated with 15 mM EDAC in PBS for 1 h at 30 °C. Cells were quickly washed twice with cold PBS without Ca^{2+} or Mg^{2+}, lysed in RIPA on ice for 15 min, centrifuged at 15,000 rpm at 4 °C for 15 min, and EGF-R was immunoprecipitated using monoclonal antibody 528. After 3–10% SDS-PAGE, receptor dimerization was quantitated by measurement of radioactivity by γ-counting of the receptor monomer-, dimer-, and background-containing excised gel bands.

RESULTS

Antibody-induced EGF-R Autophosphorylation in Cells—In a previous study we have demonstrated that in the presence of detergent, anti-EGF-R antibody 2E9, directed against the extracellular domain, and antiserum 281–7, directed against the intracellular domain of the EGF-R, were able to activate the EGF-R tyrosine kinase in membrane preparations and whole cells (30). Monoclonal antibodies 2E9, 2D11, 528, and 225 have all been reported to be unable to activate the receptor in intact cells (26, 31). To extend our previous data, we investigated whether these antibodies would also be able to induce EGF-R autophosphorylation in the presence of detergent. A431 cells were incubated with EGF or antibody in permeabilizing buffer containing 0.005% saponin and [γ-^{32}P] ATP for 10 or 20 min on ice, lysed in RIPA, and EGF-R was immunoprecipitated using antiserum 281–7. Fig. 1 shows that all antibodies tested stimulate EGF-R phosphorylation, particularly after 10 min of incubation. The stimulatory effect decreases slightly with time, suggesting that saponin treatment alone might also stimulate receptor phosphorylation. Two-dimensional phosphoamino acid analysis, shown in the *lower panel* for control, EGF- and 2E9-treated cells, demonstrates that phosphorylation was almost exclusively on tyrosine residues.

Requirement for Antibody Bivalence in Stimulation of EGF-R Autophosphorylation—The observation that all five IgG antibodies stimulated EGF-R tyrosine kinase, in spite of the

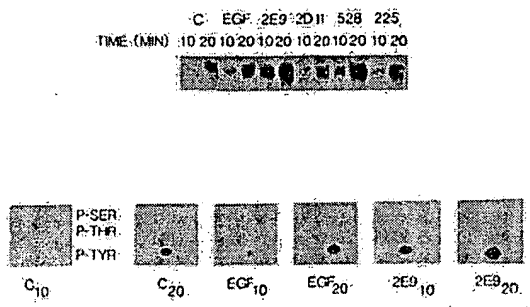


FIG. 1. **Antibody-induced EGF-R autophosphorylation in saponin-permeabilized cells.** Intact A431 cells were treated with the indicated antibodies (50 μg/ml) or EGF (50 ng/ml) in permeabilization buffer containing [γ-^{32}P]ATP and 0.005% saponin for 10 or 20 min on ice. Subsequently, the cells were lysed in RIPA, the EGF-R was immunoprecipitated, subjected to 5–15% SDS-PAGE, and autoradiographed as described under "Experimental Procedures" (*upper panel*). EGF-R-containing gel bands were excised, and protein was eluted and analyzed by two-dimensional phosphoamino acid analysis as described under "Experimental Procedures," as shown for control (*C*), EGF, and 2E9 (*lower panel*).

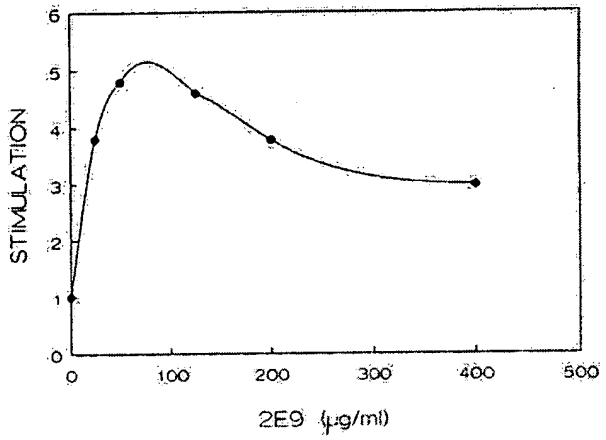


FIG. 2. **Effect of antibody 2E9 concentration on EGF-R autophosphorylation in membrane preparations.** A431 membrane preparations (20 μg) were incubated with different concentrations of 2E9 for 20 min at room temperature after which [γ-^{32}P]ATP-containing phosphorylation assay mix was added for 10 min on ice. After 5–15% SDS-PAGE and autoradiography of the gel, EGF-R-containing gel bands were excised, and radioactivity was quantitated by liquid scintillation.

fact that they are directed against different receptor epitopes, indicates that a common mechanism is involved. It has been suggested that the bivalent property of the antibodies plays an important role in receptor kinase activation (19, 30, 31). Therefore, we have studied the activation mechanism for one of the antibodies, *i.e.* 2E9, in more detail.

A431 membrane preparations were incubated with different concentrations 2E9 for 20 min at room temperature, and EGF-R phosphorylation was determined as described under "Experimental Procedures." EGF-R phosphorylation induced by 2E9 specifically occurs on tyrosine residues as determined by two-dimensional phosphoamino acid analysis (30). Maximal EGF-R autophosphorylation occurs at a narrow antibody concentration range around 50 μg/ml whereas at lower or higher antibody concentrations the level of EGF-R phosphorylation decreased (Fig. 2). This result is consistent with data obtained with other antibodies (19, 31). The decrease of EGF-R phosphorylation at high concentrations supports the assumption that antibody bivalence is indeed an important parameter in EGF-R phosphorylation since at high antibody concentrations antibody binding will be oversaturated and occurs mainly in a monovalent fashion. In contrast, at low antibody concentrations, binding will occur in a bivalent fashion but is unsaturated, and therefore the EGF-R phosphorylation is submaximal.

In order to investigate directly the importance of antibody bivalence, monovalent 2E9-Fab fragments were prepared by papain digestion and affinity purification. Subsequent SDS-PAGE revealed no intact 50-kDa heavy chain band, indicating complete digestion and purification (not shown). The Fab fragments were unable to precipitate the EGF-R in combination with protein A-Sepharose (indicating the lack of Fc fragment); however, when coupled via a rabbit anti-mouse antibody (RAM) to protein A-Sepharose, EGF-R precipitating ability was regained, demonstrating that Fab fragments were able to bind to the EGF-R. As shown in Fig. 3, the Fab fragments (100 μg/ml) were unable to stimulate EGF-R phosphorylation in A431 membrane preparations. Similar results were obtained using a Fab concentration range from 10 to 200 μg/ml (not shown). This observation is consistent with data obtained with monovalent Fab fragments of another antibody (19). However, preincubation of membranes with Fab fragments (100 μg/ml) followed by an incubation with RAM or SwAM, thereby restoring the bivalent character of the EGF-R binding complex, caused stimulation of EGF-R phosphorylation (Fig. 3).

Also in 0.005% saponin-permeabilized A431 cells, maximal EGF-R autophosphorylation occurs at an antibody concentration range around 50 μg/ml whereas at lower or higher antibody concentrations the level of EGF-R phosphorylation decreased (Fig. 4). As shown in Fig. 4, monovalent Fab fragments are also unable to induce EGF-R phosphorylation in permeabilized cells. Upon incubation with a secondary bivalent antibody, EGF-R autophosphorylation was increased (Fig. 4). The ability of the Fab-RAM binding complex to induce EGF-R phosphorylation was RAM antibody concentration dependent (optimal range around 500 μg/ml) in accordance with a role for bivalence in EGF-R activation. The data presented in Figs. 2–4 indicate that antibody bivalence

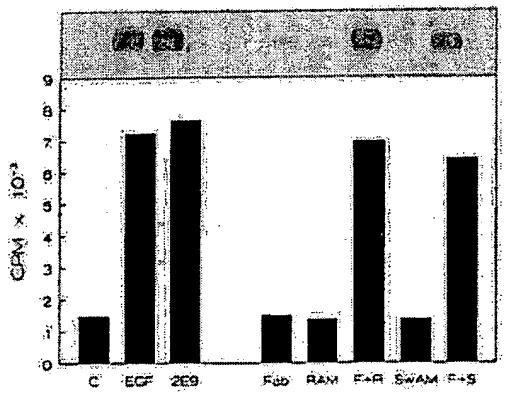


FIG. 3. **Effect of antibody bivalence on EGF-R autophosphorylation in membrane preparations.** A431 membrane preparations (20 μg) were preincubated as indicated with 100 μg/ml monovalent Fab fragments on ice for 2 h and washed by centrifugation. Subsequently, membranes were incubated as indicated for 20 min at room temperature with buffer alone (*C*), 1 μg/ml EGF, 50 μg/ml 2E9, 500 μg/ml RAM, or swine anti-mouse antiserum (*SwAM*). Fab fragment-preincubated membranes were incubated with buffer alone (*Fab*) or with 500 μg/ml RAM (*F + R*) or SwAM (*F + S*). [γ-^{32}P]ATP-containing phosphorylation assay mix was added for 10 min on ice, sample buffer was added, samples were applied to 5–15% SDS-PAGE, autoradiographed (*inset*), and radioactivity in excised gel bands was quantitated by liquid scintillation.

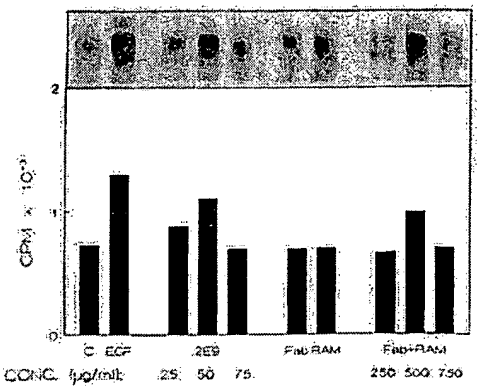


FIG. 4. Effects of antibody bivalence and concentration on EGF-R autophosphorylation in saponin-permeabilized cells. Intact A431-cells were preincubated with 100 μg/ml monovalent Fab fragments or buffer alone for 2 h on ice followed by two washes to remove unbound Fab fragments. Subsequently, untreated cells were incubated with buffer alone (*C*), 25, 50, or 75 μg/ml 2E9, or 500 μg/ml RAM. Fab fragment preincubated cells were incubated with buffer alone (*Fab*) or with 250, 500, or 750 μg/ml RAM (*Fab + RAM*). Ligand incubation was for 20 min at room temperature followed by the addition of 0.005% saponin and [γ-^{32}P]ATP-containing permeabilization buffer for 10 min on ice, eventually in the presence of 1 μg/ml EGF. EGF-R was immunoprecipitated using antibody 528 and applied to 5–15% SDS-PAGE. After autoradiography (*inset*), EGF-R-containing gel bands were excised and radioactivity measured by liquid scintillation.

is essential for increasing EGF-R autophosphorylation.

Requirement for Antibody Bivalence in EGF-R Tyrosine Kinase Activation—The observed increase of EGF-R autophosphorylation by bivalent antibody or monovalent Fab fragments in combination with a second antibody could well be the result of facilitated intermolecular cross-phosphorylation rather than tyrosine kinase activation (see "Discussion"). Therefore, we have studied the effects of these ligands on exogenous and endogenous substrate phosphorylation in membrane preparations and permeabilized cells, respectively. A431 membranes were preincubated with or without Fab fragments and subsequently incubated with ligands for 20 min at room temperature as indicated. After this incubation a phosphorylation mixture was added containing [γ-^{32}P]ATP and 2 mM angiotensin I as an exogenous substrate. Incubation of membrane preparations with EGF or 2E9 results in an increase of angiotensin I phosphorylation, demonstrating activation of the tyrosine kinase under these conditions (Fig. 5). When the membranes are incubated with the monovalent Fab fragments, no induction of tyrosine kinase activity was observed, unless followed by treatment with RAM as a second antibody. Furthermore, also in this case, tyrosine kinase activation was strongly concentration dependent for 2E9 or, in the case of Fab preincubation, RAM (Fig. 5). In accordance with data reported previously (30), 2E9 did not induce EGF-R tyrosine kinase activation in the absence of detergent (not shown).

Using the permeabilized cell system we measured endogenous substrate phosphorylation as well. It was found that several substrates could be detected upon EGF, 2E9, and Fab-RAM incubation, among which was a major 35-kDa substrate protein. Again, incubation with monovalent Fab fragments alone did not stimulate any substrate phosphorylation (Fig. 6). Upon omission of Ca^{2+} from the permeabilization buffer the 35-kDa protein was no longer observed (not shown), suggesting that this is the Ca^{2+}-dependent EGF-R substrate lipocortin I/calpactin II (35–37). These observations demonstrate that also for antibody-mediated stimulation of the

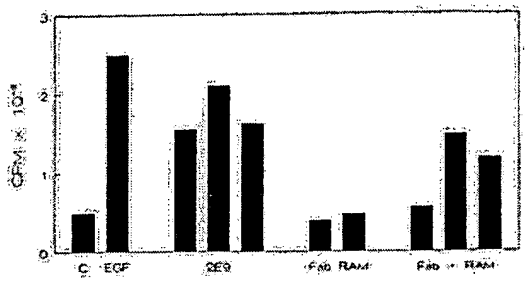


FIG. 5. Effects of antibody bivalence and concentration on EGF-R tyrosine kinase activity toward exogenous peptide in membrane preparations. Ligand incubation of A431 membranes was performed as described in the legend of Fig. 3. Subsequently, exogenous peptide angiotensin I phosphorylation was measured exactly as described under "Experimental Procedures." Incubation with buffer (*C*), EGF (1 μg/ml), 25, 50, or 75 μg/ml 2E9, Fab fragments (100 μg/ml), RAM (500 μg/ml), or 250, 500, or 750 μg/ml RAM after Fab fragment preincubation (*Fab + RAM*) are presented.

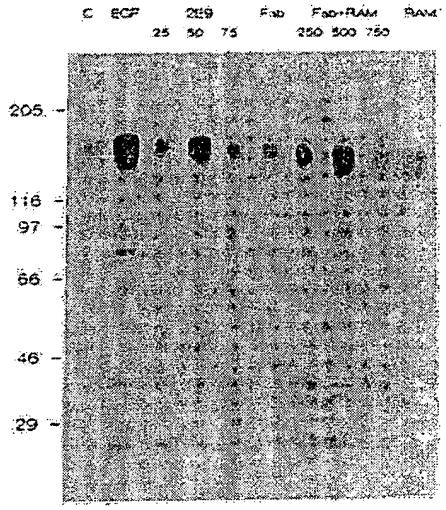


FIG. 6. Effects of antibody bivalence and concentration on EGF-R tyrosine kinase activity toward endogenous substrates in saponin-permeabilized cells. Ligand incubation and phosphorylation reactions were performed exactly as described in the legend of Fig. 4. Subsequently, cells were scraped in sample buffer, and after 5–15% SDS-PAGE and autoradiography, protein phosphorylation was analyzed by densitometric laser scanning of the autoradiograph on a Joyce-Loebl Chromoscan 3 apparatus.

EGF-R tyrosine kinase activity toward other substrates, bivalence is absolutely necessary.

Antibody-induced EGF-R Dimerization—The data presented in Figs. 3–6 strongly suggest that antibodies activate the EGF-R tyrosine kinase through induction of receptor dimerization. In order to detect possible EGF-R dimer complexes, the chemical cross-linker EDAC (1-ethyl-3-(3-dimethylaminopropyl)carbodiimide) was used, which covalently links receptor dimers. After incubation with the indicated ligands (Fig. 7) at room temperature for 20 min, phosphorylation in the presence of [γ-^{32}P]ATP was allowed to proceed for 10 min on ice and was stopped by the addition of 10 mM EDTA. Subsequently, 15 mM EDAC was added for 15 min at 30 °C. As shown in Fig. 7 (*three leftmost lanes*), antibody 2E9 induced EGF-R dimerization just as strongly as EGF. As measured by counting radioactivity in the excised EGF-R monomer- and dimer-containing gel bands, about 3% of the radioactivity was present in the EGF-R dimer fraction in untreated membranes whereas both EGF and 2E9 treatment resulted in an enhancement up to 15%.

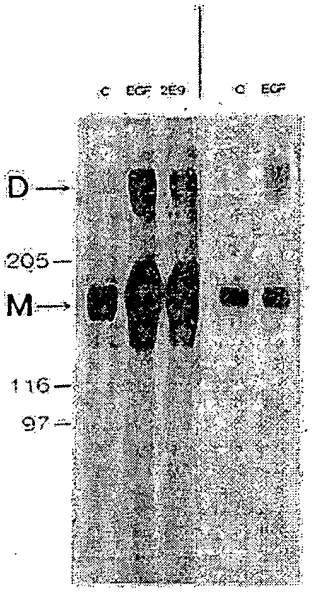


FIG. 7. Antibody-induced EGF-R dimerization in membrane preparations. A431 membranes were incubated with buffer alone (*C*), 1 μg/ml EGF, or 50 μg/ml 2E9 for 20 min at room temperature. After ligand incubation, [γ-^{32}P]ATP-containing phosphorylation assay mix was added for 10 min on ice, phosphorylation stopped by the addition of 10 mM EDTA, membranes were incubated for 15 min at 30 °C with 15 mM EDAC, and boiled in sample buffer (*three leftmost lanes*). For control experiments, after ligand incubation, 15 mM EDAC was added for 15 min at 30 °C, cross-linking was quenched by the addition of 0.5 M Tris followed by the addition of 1 μg/ml EGF and [γ-^{32}P]ATP-containing phosphorylation assay mix. Samples were boiled in sample buffer after a 10-min incubation on ice (*two rightmost lanes*). After 3–10% SDS-PAGE and autoradiography, EGF-R monomer (*M*)-, dimer (*D*)-, and background (*in between*)-containing gel bands were excised, and radioactivity was quantitated by liquid scintillation.

The possibility that receptor dimers preexist in the untreated membranes was investigated since these dimers may become phosphorylated upon ligand incubation only. Membranes were incubated with or without EGF for 20 min at room temperature, treated with EDAC for 15 min at 30 °C after which the cross-link reaction was quenched with 0.5 M Tris, and a phosphorylation reaction was carried out for 10 min on ice in the presence of EGF. As shown in Fig. 7 (*two rightmost lanes*), no EGF-R dimers were detected in the control membranes. This indicates that the observed increase in radioactivity in the dimer bands (Fig. 7, *three leftmost lanes*) was indeed due to an increase in the amount of EGF-R dimers. Furthermore, cross-linking of EGF-R dimers after EGF incubation but prior to the phosphorylation reaction resulted in enhanced EGF-induced phosphorylation compared with the control membranes (Fig. 7, *two rightmost lanes*). This suggests enhanced intermolecular cross-phosphorylation or tyrosine kinase activation in the EGF-R dimer complex. Thus, EGF-R dimerization is correlated with receptor activation. Importantly, the observation that receptor dimers are formed prior to addition of the ATP-containing phosphorylation mixture (Fig. 7, *two rightmost lanes*) indicates that dimerization also occurs in the absence of ATP and is thus not dependent on the phosphorylation reaction.

Requirement for Antibody Bivalence in Induction of EGF-R Dimerization—Next, we investigated the importance of antibody bivalence in the induction of EGF-R dimerization. Membrane preparations were incubated with the indicated ligands (Fig. 8) for 20 min at room temperature, phosphorylation

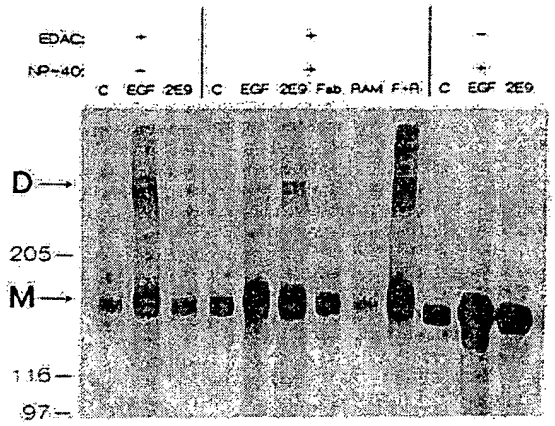


FIG. 8. Effects of detergent and antibody bivalence on EGF-R dimerization in membrane preparations. Ligand incubation of the membrane preparations was exactly as described in the legend to Fig. 3. After ligand incubation, membranes were phosphorylated by the addition of [γ-^{32}P]ATP-containing phosphorylation assay mix with (+) or without Nonidet P-40 (−) for 10 min on ice. After quenching the phosphorylation reaction by the addition of 10 mM EDTA, receptor dimers were cross-linked by the addition of 15 mM EDAC (+) or buffer alone (−) for 15 min at 30 °C. Samples were boiled in sample buffer, applied to 3–10% SDS-PAGE, and the gel was autoradiographed. Receptor dimerization was quantitated by excision of monomer (*M*)-, dimer (*D*)-, and background (*in between*)-containing gel bands and liquid scintillation of these bands.

assay mix was added for 10 min on ice, and after quenching of phosphorylation with EDTA, EDAC was added for 15 min at 30 °C. As shown in Fig. 8, no 2E9-induced receptor dimerization is observed when detergent is omitted from the phosphorylation mix. This observation confirms the disability of the anti-EGF-R antibodies to stimulate the tyrosine kinase in the absence of detergent (30) and corroborates a role for receptor dimerization in kinase activation. In agreement with this, monovalent 2E9-Fab fragments lacked the ability of inducing receptor dimerization (Fig. 8). Last but not least, also in this case, the dimer-inducing ability could be restored upon RAM incubation subsequent to the Fab fragment preincubation.

Relationship between EGF-R Dimerization and Tyrosine Kinase Activation—The data presented thus far suggest that antibody-induced EGF-R tyrosine kinase activation is the result of enhanced dimerization due to antibody bivalence. To establish further the mutual relationship between these two receptor characteristics, we have determined the effect of the specific EGF-R tyrosine kinase inhibitor tyrphostin AG 213 (a kind gift of Dr. A. Levitzki, Hebrew University, Jerusalem, Israel; Refs. 38 and 39) on EGF-induced receptor dimerization. In control experiments, using this inhibitor, we measured 80% reduction in EGF-R tyrosine kinase activation by EGF as determined by phosphoamino acid analysis of the EGF-R (not shown). As shown in Fig. 9, EGF induces EGF-R dimerization in intact surface-iodinated A431 cells. After preincubation of the A431 cells with the kinase inhibitor, EGF still induces receptor dimerization to the same extent. This result shows that receptor dimerization occurs efficiently even when tyrosine kinase activity is drastically reduced and taken together with the other data demonstrates a causative role for dimerization in activation of the EGF-R protein-tyrosine kinase and not vice versa.

DISCUSSION

In the present study we used anti-EGF-R antibodies to examine the relationship among antibody bivalence, EGF-R

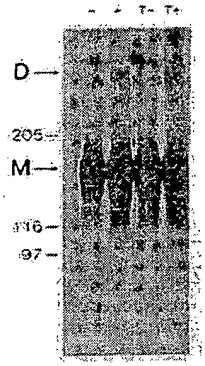


FIG. 9. EGF-induced EGF-R dimerization in tyrphostin AG 213-treated surface iodinated cells. A431 cells were grown to subconfluence, where indicated in the presence of 100 μM tyrphostin AG 213 for the last 16 h (*T*). After surface iodination with $Na^{125}I$, cells were treated for 1 h with (+) or without (−) 500 ng/ml EGF on ice. Cells were incubated with 15 mM EDAC at 30 °C for 1 h and EGF-R was immunoprecipitated with monoclonal antibody 528. After 3–10% SDS-PAGE, receptor monomer (*M*), dimer (*D*), and background (*in between*) bands were excised, and radioactivity was measured by γ-counting of these bands.

dimerization, and EGF-R tyrosine kinase activity in both membrane preparations and cells. Modification of antibody bivalence, the use of a specific EGF-R tyrosine kinase inhibitor, and the detergent dependence of antibody-induced EGF-R activation enabled us to manipulate EGF-R dimerization as well as tyrosine kinase activity. We demonstrate that (i) anti-EGF-R antibodies activate the EGF-R tyrosine kinase irrespective of the recognized epitope; (ii) antibody binding induces EGF-R dimerization; (iii) antibody-induced EGF-R dimerization as well as tyrosine kinase activation are dependent on antibody bivalence and the presence of detergent; and (iv) EGF-R dimerization is not dependent on receptor autophosphorylation and tyrosine kinase activity. These data provide strong evidence in favor of the intermolecular, active dimer, EGF-R tyrosine kinase activation model.

In initial studies, Yarden and Schlessinger (19) partly based their intermolecular activation model on data obtained with antibody-induced EGF-R autophosphorylation in Triton X-100-solubilized EGF-R. EGF-R tyrosine kinase activity was determined by EGF-R autophosphorylation, which was concluded to be an intramolecular event (19). A number of recent studies, however, indicate that receptor autophosphorylation can be mediated by intermolecular cross-phosphorylation, probably facilitated by receptor oligomerization (26, 40, 41). Since detection of enhanced EGF-R autophosphorylation does not necessarily reflect tyrosine kinase activation and to eliminate the possibility that enhanced EGF-R autophosphorylation is merely due to facilitated intermolecular cross-phosphorylation, we performed exo- and endogenous substrate phosphorylation experiments in membrane preparations and permeabilized cells. Our results show unequivocally that the EGF-R tyrosine kinase toward other substrates is indeed activated by the bivalent antibody and the combination Fab-RAM.

Concerning the role for EGF-R dimerization in tyrosine kinase activation, the question can be raised as to whether dimerization is dependent on tyrosine kinase activity. In order to address this question, we made use of the specific EGF-R tyrosine kinase inhibitor tyrphostin AG 213. The observation that incubation of A431 cells with the tyrphostin does not reduce EGF-induced formation of dimers is in agreement with the reported EGF-induced dimerization of EGF-R mutants devoid of kinase activity (42, 43) and supports a causative role for dimerization in tyrosine kinase activation.

The efficiency of several ligands in activation of receptors containing a tyrosine kinase domain has been shown to be dependent on their bivalence. Monoclonal antibodies stimulated the insulin receptor tyrosine kinase by cross-linking receptor subunit molecules (44). The *neu* protooncogene product, closely related to the EGF-R, can be converted into an oncogene by a point mutation in the transmembrane region. This mutation results in induction of receptor aggregation and an increased tyrosine kinase activity (45). Interestingly, the *neu* protooncogene-encoded kinase, but not the constitutively active oncogenic mutant, can be stimulated by agonistic antibodies but not by monovalent antibody-derived Fab fragments (46). Furthermore, activation of the platelet-derived growth factor receptor involves receptor dimerization and is dependent on platelet-derived growth factor bivalence (47–49).

Unlike platelet-derived growth factor or antibodies, EGF is a monovalent molecule that binds only one receptor (11). This raises the question as to what can be the driving force for EGF-induced receptor dimerization. We favor the hypothesis that EGF-R dimerization occurs through an EGF-induced conformational change in the extracellular domain of the EGF-R molecule, which causes enhanced affinity for other EGF receptors (50, 51). Evidence for such a conformational change has been reported recently in a study using purified EGF-R ectodomains (52). Subsequently, EGF-R tyrosine kinase activation occurs as a consequence of receptor-receptor interactions at the cytoplasmic domains. In this view the EGF-induced enhanced affinity for receptor-receptor interactions constitutes the driving force for EGF-R dimerization and subsequent tyrosine kinase activation. This mechanism bypasses the transmission of a conformational change through the transmembrane part of the receptor and is supported by the observation that an altered transmembrane region does not influence EGF-stimulated receptor dimerization or tyrosine kinase activity (53). We suggest that antibody binding induces EGF-R dimerization directly, thus bypassing the proposed EGF-induced extracellular conformational change.

This view on EGF-R activation is attractive in that it combines aspects of the intra- and intermolecular activation models. It does not, however, take into account the occurrence of a small subclass of high affinity receptors in intact cells, which played a major role in the original intermolecular activation model of Yarden and Schlessinger (19, 20), nor does it provide an explanation for the observed dominance of the high affinity subclass EGF-R in mediating signal transduction and biological responses (26, 54). Further experiments will be necessary to address these questions and to reveal the origin and molecular features of the high and low affinity EGF-R. At present, we favor the hypothesis that the high affinity subclass consists of receptors that are in a partly activated state, intermediary to the full conformational change induced by EGF binding.

Acknowledgments—We thank Dr. A. Levitzki (Dept. of Biological Chemistry, Hebrew University, Jerusalem, Israel) for kindly providing the tyrphostin AG 213.

REFERENCES

1. Ullrich, A., Coussens, L., Hayflick, J. S., Dull, T. J., Gray, A., Tam, A. W., Lee, J., Yarden, Y., Libermann, T. A., Schlessinger, J., Downward, J., Mayes, E. L. V., Whittle, N., Waterfield, M. D., and Seeburg, P. H. (1984). *Nature* 309, 418–425
2. Hunter, T., and Cooper, J. A. (1985). *Annu. Rev. Biochem.* 54, 897–930
3. Carpenter, G. (1987) *Annu. Rev. Biochem.* 56, 881–914

4. Prywes, R., Livneh, E., Ullrich, A., and Schlessinger, J. (1986). *EMBO J.* **5**, 2179–2190
5. Livneh, E., Prywes, R., Kashles, O., Reiss, N., Sasson, I., Mory, Y., Ullrich, A., and Schlessinger, J. (1986) *J. Biol. Chem.* **261**, 12490–12497
6. Chen, W. S., Lazar, C. S., Poenie, M., Tsien, R. Y., Gill, G. N., and Rosenfeld, M. G. (1987) *Nature* **328**, 820–823
7. Honegger, A. M., Dull, T. J., Felder, S., van Obberghen, E., Bellot, F., Szapary, D., Schmidt, A., Ullrich, A., and Schlessinger, J. (1987) *Cell* **51**, 199–209
8. Honegger, A. M., Szapary, D., Schmidt, A., Lyall, R., van Obberghen, E., Dull, T. J., Ullrich, A., and Schlessinger, J. (1987) *Mol. Cell. Biol.* **7**, 4568–4571
9. Moolenaar, W. H., Bierman, A. J., Tilly, B. C., Verlaan, I., Defize, L. H. K., Honegger, A. M., Ullrich, A., and Schlessinger, J. (1988) *EMBO J.* **7**, 707–710
10. Bertics, P. J., and Gill, G. N. (1985) *J. Biol. Chem.* **260**, 14642–14647
11. Weber, W., Bertics, P. J., and Gill, G. N. (1984) *J. Biol. Chem.* **259**, 14631–14636
12. Gill, G. N., Bertics, P. J., and Santon, J. B. (1987) *Mol. Cell. Endocrinol.* **51**, 169–186
13. Koland, J. G., and Cerione, R. A. (1988) *J. Biol. Chem.* **263**, 2230–2237
14. Schlessinger, J. (1986) *J. Cell Biol.* **103**, 2067–2072
15. Schlessinger, J. (1988) *Trends Biochem. Sci.* **13**, 443–447
16. Basu, M., Sen-Majumdar, A., Basu, A., Murthy, U., and Das, M. (1986) *J. Biol. Chem.* **261**, 12879–12882
17. Biswas, R., Basu, M., Sen-Majumdar, A., and Das, M. (1985) *Biochemistry* **24**, 3795–3802
18. Basu, A., Raghunath, M., Bishayee, S., and Das, M. (1989) *Mol. Cell. Biol.* **9**, 671–677
19. Yarden, Y., and Schlessinger, J. (1987) *Biochemistry* **26**, 1434–1442
20. Yarden, Y., and Schlessinger, J. (1987) *Biochemistry* **26**, 1443–1451
21. Boni-Schnetzler, M., and Pilch, P. F. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84**, 7832–7836
22. Fanger, B. O., Stephens, J. E., and Staros, J. V. (1989) *FASEB J.* **3**, 71–75
23. Northwood, I. C., and Davis, R. J. (1988) *J. Biol. Chem.* **263**, 7450–7453
24. Fanger, B. O., Austin, K. S., Earp, H. S., and Cidlowski, J. A. (1986) *Biochemistry* **25**, 6414–6420
25. Cochet, C., Kashles, O., Chambaz, E. M., Borello, I., King, C. R., and Schlessinger, J. (1989) *J. Biol. Chem.* **263**, 3290–3295
26. Defize, L. H. K., Boonstra, J., Meisenhelder, J., Kruijer, W., Tertoolen, L. G. J., Tilly, B. C., Hunter, T., van Bergen en Henegouwen, P. M. P., Moolenaar, W. H., and de Laat, S. W. (1989) *J. Cell Biol.* **109**, 2495–2507
27. Schreiber, A. B., Lax, I., Yarden, Y., Eshhar, Z., and Schlessinger, J. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78**, 7535–7539
28. Schreiber, A. B., Libermann, T. A., Lax, I., Yarden, Y., and Schlessinger, J. (1983) *J. Biol. Chem.* **258**, 846–853
29. Defize, L. H. K., Moolenaar, W. H., van der Saag, P. T., and de Laat, S. W. (1986) *EMBO J.* **5**, 1187–1192
30. Spaargaren, M., Defize, L. H. K., de Laat, S. W., and Boonstra, J. (1990) *Biochem. Biophys. Res. Commun.* **171**, 882–889
31. Gill, G. N., Kawamoto, T., Cochet, C., Le, A., Sato, J. D., Masui, H., McLeod, C., and Mendelsohn, J. (1984) *J. Biol. Chem.* **259**, 7755–7760
32. Defize, L. H. K., Mummery, C. L., Moolenaar, W. H., and de Laat, S. W. (1987) *Cell Differ.* **20**, 87–102
33. Cohen, S., Ushiro, H., Stoschek, C., and Chinkers, M. (1982) *J. Biol. Chem.* **257**, 1523–1531
34. Defize, L. H. K., Arndt-Jovin, D. J., Jovin, T. M., Boonstra, J., Meisenhelder, J., Hunter, T., de Hey, H. T., and de Laat, S. W. (1988) *J. Cell Biol.* **107**, 939–949
35. Guigni, T. D., James, L. C., and Haigler, H. T. (1985) *J. Biol. Chem.* **260**, 15081–15090
36. Fava, R. A., and Cohen, S. (1984) *J. Biol. Chem.* **259**, 2636–2645
37. Pepinsky, R. B., and Sinclair, L. K. (1986) *Nature* **321**, 81–84
38. Yaish, P., Gazit, A., Gilon, C., and Levitzki, A. (1988) *Science* **242**, 933–935
39. Lyall, M., Zilberstein, A., Gazit, A., Gilon, C., Levitzki, A., and Schlessinger, J. (1989) *J. Biol. Chem.* **264**, 14503–14509
40. Honegger, A. M., Kris, R. M., Ullrich, A., and Schlessinger, J. (1989) *Proc. Natl. Acad. Sci. U. S. A.* **86**, 925–929
41. Honegger, A. M., Schmidt, A., Ullrich, A., and Schlessinger, J. (1990) *Mol. Cell. Biol.* **10**, 4035–4044
42. Glenney, J. R., Jr., Chen, W. S., Lazar, C. S., Walton, G. M., Zokas, L. M., Rosenfeld, M. G., and Gill, G. N. (1988) *Cell* **52**, 675–684
43. Margolis, B., Bellot, F., Honegger, A. M., Ullrich, A., Schlessinger, J., and Zilberstein, A. (1990) *Mol. Cell. Biol.* **10**, 435–441
44. O'Brien, R. M., Soos, M. A., and Siddle, K. (1987) *EMBO J.* **6**, 4003–4010
45. Weiner, D. B., Liu, J., Cohen, J. A., Williams, W. V., and Greene, M. I. (1989) *Nature* **339**, 230–231
46. Yarden, Y. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87**, 2569–2573
47. Bishayee, S., Majumdar, S., Khire, J., and Das, M. (1989) *J. Biol. Chem.* **264**, 11699–11705
48. Hammacher, A., Mellstrom, K., Heldin, C.-H., and Westermark, B. (1989) *EMBO J.* **8**, 2489–2495
49. Heldin, C.-H., Ernlund, A., Rorsman, C., and Rönnstrand, L. (1989) *J. Biol. Chem.* **264**, 8905–8912
50. Defize, L. H. K. (1988) *Activation of the Human Epidermal Growth Factor Receptor*, Ph.D. thesis, University of Utrecht
51. Ullrich, A., and Schlessinger, J. (1990) *Cell* **61**, 203–212
52. Greenfield, C., Hiles, I., Waterfield, M. D., Federwisch, M., Wollmer, A., Blundell, T. L., and McDonald, N. (1989) *EMBO J.* **8**, 4115–4123
53. Kashles, O., Szapary, D., Bellot, F., Ullrich, A., Schlessinger, J., and Schmidt, A. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85**, 9567–9571
54. Bellot, F., Moolenaar, W. H., Kris, R., Mirakhur, B., Verlaan, I., Ullrich, A., Schlessinger, J., and Felder, S. (1990) *J. Cell Biol.* **110**, 491–502

EXHIBIT B

Proc. Natl. Acad. Sci. USA
Vol. 89, pp. 2140–2144, March 1992
Cell Biology

Homodimerization and constitutive activation of the erythropoietin receptor

STEPHANIE S. WATOWICH*, AKIHIKO YOSHIMURA*, GREGORY D. LONGMORE*†, DOUGLAS J. HILTON*, YUKO YOSHIMURA*, AND HARVEY F. LODISH*‡

*Whitehead Institute for Biomedical Research, Nine Cambridge Center, Cambridge, MA 02142; †Division of Hematology-Oncology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115; and ‡Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139

Contributed by Harvey F. Lodish, December 17, 1991

ABSTRACT **The erythropoietin receptor (EPO-R) is a member of the recently described cytokine receptor superfamily. A constitutively active (hormone independent) form of the EPO-R was isolated that has a single amino acid change in the exoplasmic domain, converting arginine-129 to cysteine (R129C). Since EPO-Rs containing R129S, R129E, and R129P mutations are functionally wild type, the presence of cysteine at residue 129, and not the loss of arginine, is required for constitutive activity. Several mutant forms of the EPO-R were analyzed; all constitutively active mutants form disulfide-linked homodimers, whereas EPO-responsive or inactive forms of the receptor do not. Monomers and disulfide-linked dimers of the constitutive receptor are present on the plasma membrane and bind EPO with a single affinity. Homodimerization of the EPO-R is likely to play a role in ligand-induced signal transduction, and disulfide-linked dimerization of the constitutive receptor may mimic this step.**

Erythropoietin (EPO) is a serum glycoprotein hormone required for the survival, proliferation, and differentiation of committed erythroid progenitor cells. The murine EPO receptor (EPO-R) cDNA was isolated by expression cloning (1) and was found to have sequence homology with other cytokine receptors (2). Conserved structural features of the cytokine receptor superfamily include four similarly spaced exoplasmic cysteine residues, as well as a motif, WSXWS, located in the exoplasmic domain close to the membrane-spanning region (3). The EPO-R and other members of the cytokine receptor family do not contain kinase-related or nucleotide-binding consensus sequences in their cytoplasmic domains and the intracellular signaling pathways they initiate after ligand binding have yet to be defined.

Although little is known of the mechanisms by which cytokine receptors transduce their signal, dimerization of the receptors is thought to play a role. The receptors for interleukins 2, 3, 5, and 6, as well as granulocyte–macrophage colony-stimulating factor, contain at least two different subunits (4–8), while the ligand binding subunits of the granulocyte colony-stimulating factor receptor, prolactin receptor, and growth hormone receptor form homodimers (9–11). Dimerization has been postulated to yield high-affinity receptors and also to provide the first step in the signal transduction pathway (11, 12).

Expression of the cloned EPO-R cDNA in the interleukin 3-dependent pro-B-cell line BA/F3 allows the cells to grow in response to EPO, demonstrating that the EPO-R can functionally transmit a growth signal (13). The recent demonstration that the mutation of arginine-129 to cysteine (R129C) results in a constitutively active (14) and oncogenic form (15) of the EPO-R is provocative in that it implicates the formation of aberrant inter- or intramolecular disulfide bonds in the process of receptor activation.

The role of the new cysteine residue in the constitutively active receptor and the possibility that this receptor may have an altered disulfide-bonding pattern were investigated by both biochemical and mutagenesis approaches. The presence of cysteine at residue 129 is required for EPO-independent signaling. Analysis of several mutants of the EPO-R has revealed that all constitutively active mutants, but not the wild-type receptor or EPO-dependent mutants, form disulfide-linked homodimers in the endoplasmic reticulum (ER) and a fraction of these dimers are transported to the plasma membrane.

MATERIALS AND METHODS

Mutagenesis Techniques. The constitutive mutant of the EPO-R (R129C), the truncated form of the receptor (tEPO-R), and the constitutive, truncated form of the receptor (tEPO-R/R129C) were isolated by a retroviral transduction system (14). The remaining mutant EPO-Rs (see Fig. 1) were generated by polymerase chain reaction, with synthetic oligonucleotide primers encoding the desired amino acid substitutions. Mutant EPO-R cDNAs were subcloned into the mammalian expression vector pXM (16) and into M13mp18 and M13mp19 vectors. Sequences of the mutant cDNAs were confirmed by the dideoxynucleotide chain-terminating method, using synthetic oligonucleotides as primers.

Cell Culture Conditions and Transfections. The wild-type and mutant EPO-R cDNAs in pXM were introduced into BA/F3 cells by electroporation, and stable transformants were cloned as described (14). Interleukin 3-dependent, EPO-dependent, and factor-independent clones of BA/F3 cells were maintained as described (14).

Metabolic Labeling and Immunoprecipitation. BA/F3 cell lines expressing wild-type or mutant EPO-Rs were metabolically labeled with [^{35}S]methionine and cysteine (^{35}S-Express; NEN). Cell lysates were prepared in buffer containing 1% Triton X-100, 0.5% sodium deoxycholate, 0.1% SDS, 150 mM NaCl, 50 mM Tris·HCl (pH 7.4), 200 mM iodoacetamide, and 2 mM phenylmethylsulfonyl fluoride. Lysates were incubated with anti-peptide antibodies raised against N- or C-terminal peptides of the EPO-R (17), followed by incubation with protein A-agarose beads (Boehringer Mannheim). Proteins were eluted in gel sample buffer [1% SDS/10% (vol/vol) glycerol/80 mM Tris·HCl, pH 6.8] with or without 1% 2-mercaptoethanol.

Gel Electrophoresis and Immunoblot Analysis. One- and two-dimensional gel electrophoresis was carried out on SDS/7.5% polyacrylamide gels (18). For two-dimensional gels, polypeptides were separated first under nonreducing condi-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: EPO, erythropoietin; EPO-R, EPO receptor; ER, endoplasmic reticulum; t-EPO-R, truncated form of EPO-R; bEPO, biotinylated EPO.

tions, and then after reduction with 5% 2-mercaptoethanol. Immunoblotting was performed as described (19). The filters were incubated with antisera raised to EPO-R peptides, followed by ^{125}I-labeled protein A (Amersham).

Scatchard Analysis and Biotinylated-EPO Binding. EPO was iodinated by the iodine monochloride method and had a specific activity of 4×10^6 cpm/pmol. Analyses of saturation binding were performed as described (20). Biotinylated EPO (bEPO) was prepared (21) and surface EPO-Rs were isolated as described, except that cross-linking of the bEPO–receptor complex was not performed (22).

RESULTS

The substitution R129C introduces a sixth cysteine residue into the EPO-R exoplasmic domain and confers constitutive (hormone independent) activity (Fig. 1, mutant R129C; ref. 14). Mutation of residue 129 to serine, glutamic acid, or proline yields a wild-type, not a constitutive, phenotype (data not shown). Since a cysteine residue at position 129, rather than the loss of an arginine, was crucial for constitutive activity, we suspected that the new cysteine may form an intramolecular disulfide bond, possibly with C179. To test this hypothesis, C179 was mutated to a serine residue either in the wild-type receptor (C179S) or in the constitutive mutant (R129C/C179S). Neither the EPO responsiveness of the wild-type receptor nor the constitutive activity of the R129C receptor was affected by this mutation (Fig. 1), demonstrating that C179 is not involved in EPO-induced activation of the wild-type receptor or in constitutive activation of the R129C mutant.

The possibility that R129C has other intramolecular disulfide bond rearrangements or has formed other intermolecular disulfide bonds remained. To determine whether the pattern of intramolecular or intermolecular disulfide bonding in R129C was different from that in the wild-type receptor, we immunoprecipitated metabolically labeled receptors and analyzed them by reducing and nonreducing SDS/PAGE. Both the newly synthesized wild-type EPO-R and R129C migrate with an apparent molecular mass of 64 kDa under reducing conditions (data not shown). Under nonreducing conditions, the newly synthesized wild-type EPO-R migrates with an apparent molecular mass of 64 kDa (Fig. 2, lane 4), while the constitutive receptor migrates as a monomer of ≈64 kDa as well as an oligomer of ≈160 kDa (lane 7). This oligomeric species accumulates during the 2-h chase at 18°C (lane 9) and it appears to be a disulfide-linked complex of the R129C receptors since it is not visible after reduction of the samples before SDS/PAGE (data not shown).

We assayed several mutants of the EPO-R for their ability to form disulfide-linked oligomers and found that all constitutively active mutants of the EPO-R form disulfide-linked oligomers, while all hormone-responsive or inactive forms of the receptor do not (Figs. 1 and 2). The constitutive R129C/C179S mutant forms disulfide-linked oligomers, although somewhat less efficiently than R129C (Fig. 2, lanes 13–15). tEPO-R/R129C also forms disulfide-linked oligomers that migrate faster than the R129C oligomers, as expected for an oligomeric species composed of truncated receptor molecules (lanes 19–21). tEPO-R and the C179S mutant, both of which are EPO responsive like the wild-type EPO-R, do not form detectable disulfide-linked oligomers (lanes 10–12 and 16–18).

To test further the correlation found between the presence of C129 in the EPO-R, disulfide-linked oligomerization, and constitutive activation, we assayed two inactive forms of the EPO-R (ws1 and ws1/R129C) for their ability to oligomerize. Both ws1 and ws1/R129C lack three residues, AWS, from the conserved WSXWS region and have a GA substitution in their place. In addition, the ws1/R129C mutant contains the R129C mutation. When expressed in BA/F3 cells, these mutant receptors are unable to transmit an EPO growth signal since the cells will only grow in the presence of interleukin 3 (Fig. 1). Both mutants are retained in the ER after synthesis and are likely to be misfolded (data not shown). Neither mutant receptor formed disulfide-linked oligomers (Fig. 2, lanes 22–27).

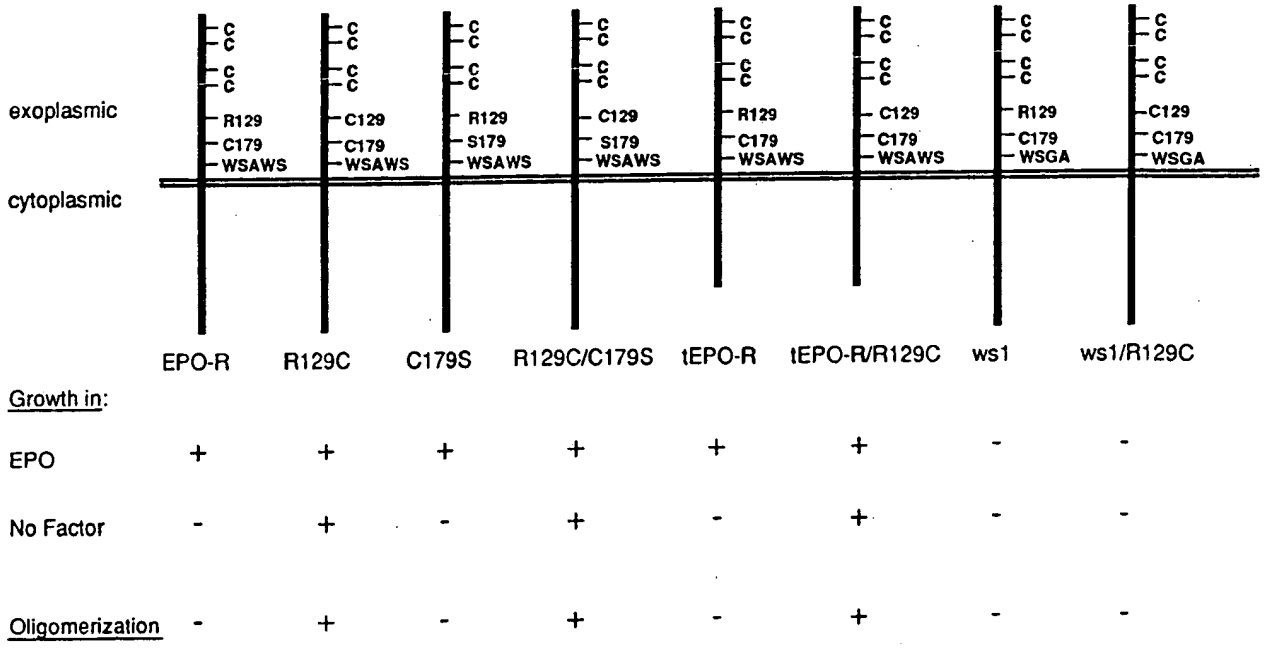


FIG. 1. Schematic diagram of mutant EPO-Rs. Mutants of the EPO-R cDNA were generated or isolated as described. The receptors were assayed for their ability to confer EPO-dependent or factor-independent growth in BA/F3 cells as described (13, 14). Assays for oligomerization were performed as indicated in the legend to Fig. 2. Residues 129 and 179 are indicated; the four conserved exoplasmic cysteine residues (C27, C37, C65, and C81) and the conserved WSAWS sequence (residues 207–211) are not numbered.

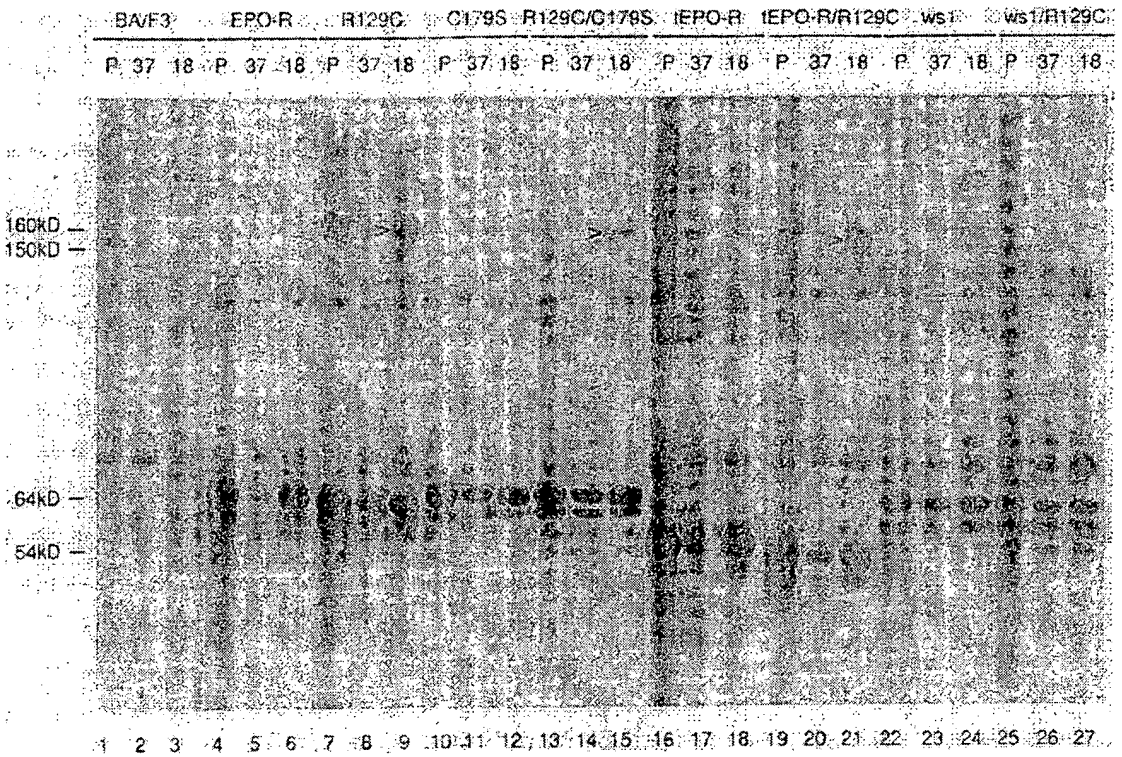


FIG. 2. Constitutively active mutants of the EPO-R form disulfide-linked oligomers. Parental BA/F3 cells or cells expressing wild-type or mutant EPO-Rs were pulse labeled (lanes P) with [^{35}S]methionine/cysteine for 10 min at 37°C and then chased for 2 h at 37°C (lanes 37) or 18°C (lanes 18). EPO-Rs were immunoprecipitated and separated by nonreducing gel electrophoresis. The migration positions of monomeric EPO-R (≈64 kDa) and monomeric, tEPO-R (≈54 kDa) are indicated on the left. The positions of oligomeric, constitutive EPO-R (≈160 kDa) and oligomeric, constitutive, tEPO-R (≈150 kDa) are indicated with arrowheads.

The identity of the polypeptides present in the EPO-R immunoprecipitates from parental BA/F3 cells and cells expressing the wild-type receptor, R129C, and tEPO-R/R129C, was determined by two-dimensional gel electrophoresis. The only polypeptides that were generated by reduction of the disulfide-linked oligomers were monomers of R129C (Fig. 3*C*) and tEPO-R/R129C (Fig. 3*D*), demonstrating that the disulfide-linked species are EPO-R homooligomers and eliminating the possibility that the constitutive EPO-R forms a heterooligomer with another polypeptide of similar size. The ability of R129C/C179S, but not the wild-type receptor or C179S, to form disulfide-linked oligomers (Fig. 2) suggests that only C129 is available for forming interchain disulfide bonds; thus, it is likely that the R129C oligomers are homodimers. The wild-type EPO-R migrated identically before and after reduction (Fig. 3*B*).

To determine whether the disulfide-linked homodimers are present on the cell surface, we used bEPO to isolate cell-surface receptors (22). Both R129C monomers and disulfide-linked dimers were found on the plasma membrane (Fig. 4, lane 3). Similarly, cells synthesizing R129C/C179S (lane 5) and tEPO-R/R129C (data not shown) expressed both monomers and dimers on the surface. After reduction, oligomers of the constitutive receptor were not detectable and only monomers were seen (lanes 8 and 10). Cells synthesizing hormone-responsive forms of the receptor (wild-type EPO-R, C179S, tEPO-R) did not express detectable levels of surface disulfide-linked dimers; as expected, only monomeric species are found on the plasma membrane (lanes 2 and 4; data not shown).

The presence of both monomers and dimers of R129C on the plasma membrane suggested that there may be two classes of surface receptors, perhaps corresponding to low- and high-affinity EPO-binding species. Scatchard analysis of ^{125}I-labeled EPO binding to BA/F3 cells expressing R129C demonstrates, however, that the surface receptors display a single affinity for EPO (K_d = 700 pM). Approximately 1000 surface receptors are expressed per cell (Fig. 5). Similarly, when EPO binding to the erythroid cell line HCD57 expressing R129C was examined, only a single class of receptors was detected even though both monomeric and dimeric forms of R129C were detected on the plasma membrane (data not shown). BA/F3 cells synthesizing wild-type EPO-R also displayed a single class of receptors (23).

DISCUSSION

The EPO-R can be activated by two different, hormone-independent mechanisms: by interaction with the gp55 glycoprotein of spleen focus-forming virus (13) and by a point mutation, R129C, in the extracellular domain (14). The presence of a cysteine residue at position 129, and not the loss of an arginine, appears to be required for constitutive activity since substitution of arginine at position 129 with serine, proline, or glutamic acid does not alter the ability of the EPO-R to confer EPO-responsive growth in BA/F3 cells. This requirement for a cysteine residue led to the hypothesis that the constitutive receptor may form intramolecular or intermolecular disulfide bonds, which alter the conformation of the receptor and render it constitutively active.

By analogy with the intramolecular disulfide bonding pattern of the growth hormone receptor (24), it could be predicted that the first and second cysteine residues (C27 and C37) of the EPO-R form a disulfide bond as do the third and fourth (C65 and C81), leaving C179 unpaired. Since the C179S mutant is functionally wild-type, C179 is not essential for normal receptor function, including ligand binding. The phenotype of mutant R129C/C179S demonstrates that C179

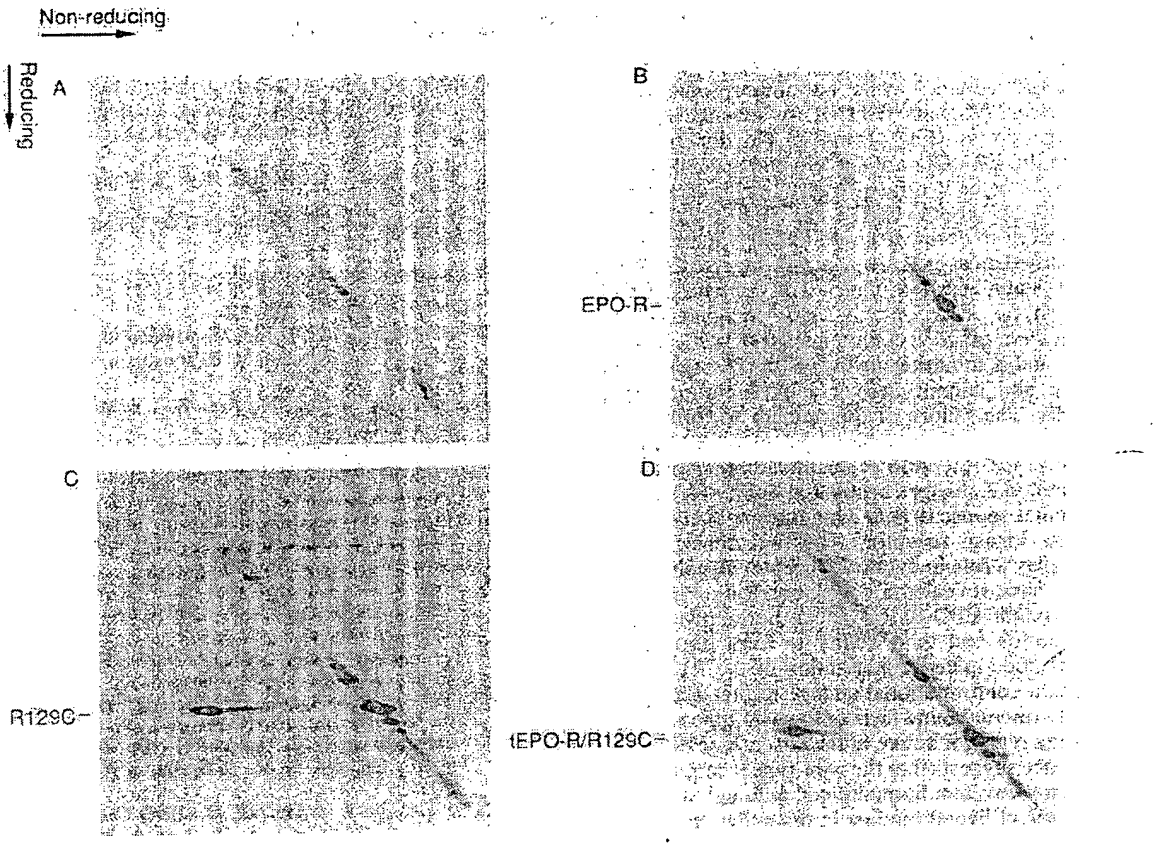


FIG. 3. R129C and tEPO-R/R129C form disulfide-linked homodimers. Parental BA/F3 cells (*A*) or cells expressing the wild-type EPO-R (*B*), R129C (*C*), or tEPO-R/R129C (*D*) were pulse labeled with [^{35}S]methionine/cysteine for 10 min at 37°C and then chased for 2 h at 18°C. The EPO-Rs were immunoprecipitated and separated under nonreducing conditions in the first dimension, followed by reduction and electrophoresis in the second dimension. The positions of monomeric EPO-Rs are indicated.

is also not required for constitutive activity of the receptor; therefore, in the R129C mutant, C179 is unlikely to pair with C129 (Fig. 1). Since the constitutively active R129C receptor binds EPO with an affinity similar to that of the wild-type receptor (Fig. 5), it is unlikely that intramolecular disulfide bond rearrangements have occurred. We conclude that R129C and R129C/C179S form disulfide-linked homodimers through C129.

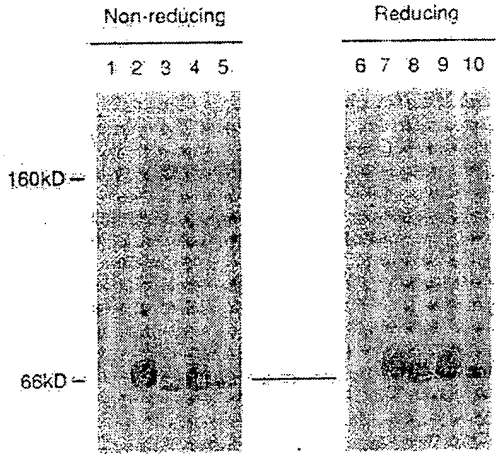


FIG. 4. Disulfide-linked dimers of the constitutive EPO-R are found on the cell surface. Parental BA/F3 cells (lanes 1 and 6) or cells expressing the wild-type EPO-R (lanes 2 and 7), R129C (lanes 3 and 8), C179S (lanes 4 and 9), or R129C/C179S (lanes 5 and 10) were incubated with 10 nM bEPO for 4–6 h at 4°C. The cells were washed and then lysed in buffer containing 0.5% Nonidet P-40 and 200 mM iodoacetamide. Surface receptors that had bound bEPO were isolated on streptavidin-agarose beads, separated by nonreducing and reducing gel electrophoresis, and assayed by immunoblotting. The positions of cell-surface monomeric EPO-Rs (66 kDa) and surface disulfide-linked dimers (160 kDa) are indicated.

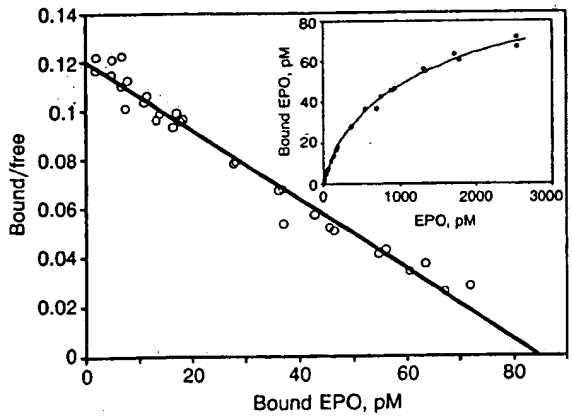


FIG. 5. EPO binding to BA/F3 cells expressing R129C. Approximately 5 × 10^6 cells expressing R129C were incubated with various concentrations of ^{125}I-labeled EPO, in the presence and absence of 60 nM unlabeled EPO, in a vol of 100 μl for 14 h at 4°C. Free ^{125}I-labeled EPO was separated from the bound hormone, and specific binding was determined as described (20). A saturation curve (*Inset*) and a Scatchard plot of the data are shown.

The formation of disulfide-linked dimers is correlated with constitutive activity. All constitutively active mutants of the EPO-R (R129C, tEPO-R/R129C, and R129C/C179S) form disulfide-linked dimers, while hormone-responsive (wild-type EPO-R, C179S, and tEPO-R) or inactive (ws1, ws1/R129C) forms do not (Figs. 1 and 2). The dimers appear to assemble in the ER (data not shown) and a small proportion of them reach the cell surface, where they bind EPO with a single affinity of 700 pM (Figs. 4 and 5). The R129C/C179S mutant appears to form disulfide-linked dimers somewhat less efficiently than does R129C, suggesting that C179 may be involved in stabilizing the dimer. The presence of C129 is necessary but not sufficient for oligomerization and constitutive activation since the mutant ws1/R129C, containing both C129 and a mutation in the WSXWS region, fails to oligomerize and cannot deliver a proliferation signal in BA/F3 cells. The failure of this mutant to form disulfide-linked dimers also indicates that covalent dimerization is not simply due to ER retention of misfolded receptors.

Conformational changes brought about by receptor oligomerization in response to ligand binding are likely to activate the tyrosine kinase receptors (25) and dimerization also appears to play a role in cytokine receptor signaling (11, 12). Our results have revealed a strong correlation between the ability of mutant EPO-Rs to induce hormone-independent cell proliferation and their ability to form disulfide-linked dimers. Disulfide-linked dimerization of the R129C mutant may induce a conformational change in the receptor, mimicking the hormone-bound form of the wild-type receptor and rendering the receptor active in the absence of EPO. Preliminary evidence suggests that the wild-type receptor is capable of forming noncovalent homodimers, although we do not yet know the role of ligand binding in dimerization.

C129 is likely to be present at the dimer interface of the disulfide-linked R129C receptors and by extrapolation R129 or neighboring residues may play a role in the noncovalent dimerization of the wild-type receptor. Dimerization of the EPO-R may be analogous to the growth hormone receptor, where it has been shown that two receptor subunits bind to different sites on a single growth hormone molecule, and the receptor dimers are stabilized by growth hormone (11). Additional experiments could be directed toward elucidating the structural features involved in EPO-R dimerization and understanding the role of dimerization in receptor activation.

Note Added in Proof. The crystal structure of the growth hormone receptor–growth hormone complex has recently been determined (26). Sequence comparison between the EPO-R and the growth hormone receptor has revealed that residue 129 of the EPO-R would fall in the region corresponding to the growth hormone receptor dimer interface. These observations support the hypothesis that EPO may signal through formation of a noncovalent homodimer of the EPO-R.

We thank Dr. Drorit Neumann for many helpful discussions and for critically reading the manuscript. We also thank Dr. K. Shimosaka and Mr. T. Hirano (Kirin Brewery, Yokohama, Japan) for their generous gift of pure, recombinant EPO. This work was supported by Fellowship 710 from the Anna Fuller Fund and Fellowship PF-3546 from the American Cancer Society to S.S.W., a fellowship from the Nakatomi Health Science Foundation (Saga, Japan) to A.Y., National Institutes of Health/National Institute on Aging Physician Scientist Award AG00294 to G.D.L., a postdoctoral fellowship from the Lucille P. Markey charitable trust to D.J.H., and National Institutes of Health Grant HL32262 to H.F.L.

1. D'Andrea, A. D., Lodish, H. F. & Wong, G. G. (1989) *Cell* **57**, 277–285.
2. D'Andrea, A. D., Fasman, G. D. & Lodish, H. F. (1989) *Cell* **58**, 1023–1024.
3. Cosman, D., Lyman, S. D., Idzerda, R. L., Beckmann, M. P., Park, L. S., Goodwin, R. G. & March, C. J. (1990) *Trends Biochem. Sci.* **15**, 265–270.
4. Hatakeyama, M., Tsudo, M., Minamoto, S., Kono, T., Doi, T., Miyata, T., Miyasaka, M. & Taniguchi, T. (1989) *Science* **244**, 551–556.
5. Kitamura, T., Sato, N., Arai, K.-I. & Miyajima, A. (1991) *Cell* **66**, 1165–1174.
6. Devos, R., Plaetinck, G., Van der Heyden, J., Cornelis, S., Vandekerckhove, J., Fiers, W. & Tavernier, J. (1991) *EMBO J.* **10**, 2133–2137.
7. Taga, T., Hibi, M., Hirata, Y., Yamasaki, K., Yasukawa, K., Matsuda, T., Hirano, T. & Kishimoto, T. (1989) *Cell* **58**, 573–581.
8. Hayashida, K., Kitamura, T., Gorman, D. M., Arai, K.-I., Yokota, T. & Miyajima, A. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 9655–9659.
9. Fukunaga, R., Ishizaka-Ikeda, E. & Nagata, S. (1990) *J. Biol. Chem.* **265**, 14008–14015.
10. Rozakis-Adcock, M. & Kelly, P. A. (1991) *J. Biol. Chem.* **266**, 16472–16477.
11. Cunningham, B. C., Ultsch, M., De Vos, A. M., Mulkerrin, M. G., Clauser, K. R. & Wells, J. A. (1991) *Science* **254**, 821–825.
12. Nicola, N. A. & Metcalf, D. (1991) *Cell* **67**, 1–4.
13. Li, J.-P., D'Andrea, A. D., Lodish, H. F. & Baltimore, D. (1990) *Nature (London)* **343**, 762–764.
14. Yoshimura, A., Longmore, G. & Lodish, H. F. (1990) *Nature (London)* **348**, 647–649.
15. Longmore, G. D. & Lodish, H. F. (1991) *Cell* **67**, 1089–1102.
16. Yang, Y.-C., Ciarletta, A. B., Temple, P. A., Chung, M. P., Kovacic, S., Witek-Giannotti, J. S., Leary, A. C., Kriz, R., Donahue, R. E., Wong, G. G. & Clark, S. C. (1986) *Cell* **47**, 3–10.
17. Yoshimura, A., D'Andrea, A. D. & Lodish, H. F. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 4139–4143.
18. Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685.
19. Burnette, W. N. (1981) *Anal. Biochem.* **112**, 195–203.
20. Hilton, D. J., Nicola, N. A. & Metcalf, D. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 5971–5975.
21. Wognum, A. W., Lansdorp, P. M., Humphries, R. K. & Krystal, G. (1990) *Blood* **76**, 697–705.
22. Yoshimura, A. & Lodish, H. F. (1992) *Mol. Cell. Biol.* **12**, 706–715.
23. D'Andrea, A. D., Yoshimura, A., Youssoufian, H., Zon, L. I., Koo, J.-W. & Lodish, H. F. (1991) *Mol. Cell. Biol.* **11**, 1980–1987.
24. Fuh, G., Mulkerrin, M. G., Bass, S., McFarland, N., Brochier, M., Bourell, J. H., Light, D. R. & Wells, J. A. (1990) *J. Biol. Chem.* **265**, 3111–3115.
25. Ullrich, A. & Schlessinger, J. (1990) *Cell* **61**, 203–212.
26. De Vos, A. M., Ultsch, M. & Kossiakoff, A. A. (1992) *Science* **255**, 306–312.

11. G. M. Clore *et al.*, *EMBO J.* 5, 2729 (1986).
12. Upper distance limits for distances involving methyl and methylene protons were corrected appropriately for center averaging [K. Wüthrich, M. Billeter, W. Braun, *J. Mol. Biol.* 169, 949 (1983)]. In addition, 0.5 Å was added to the upper limits of distances involving methyl protons to account for the higher apparent intensity of methyl resonances [G. M. Clore, A. M. Gronenborn, M. Nilges, C. A. Ryan, *Biochemistry* 26, 8012 (1987)].
13. The $^3J_{HNH\alpha}$ coupling constants were obtained from an HMQC *J* spectrum [L. E. Kay and A. Bax, *J. Magn. Reson.* 86, 110 (1990)] as described previously [J. D. Forman-Kay, A. M. Gronenborn, L. E. Kay, P. T. Wingfield, G. M. Clore, *Biochemistry* 29, 1566 (1990)]. For $^3J_{HN\alpha} < 6$ Hz and > 8 Hz, ϕ was restrained to $-50 \pm 40°$ and $-125 \pm 50°$, respectively [A. Pardi, M. Billeter, K. Wüthrich, *J. Mol. Biol.* 180, 741 (1984)].
14. Restraints on ϕ of $-50 \pm 60°$ in the regions of regular helix, as delineated by NH(*i*)-NH(*i*+1), CαH(*i*)-NH(*i*+2,3,4) and CαH(*i*)-CβH(*i*+3) NOEs and the presence of slowly exchanging amide protons (*16*), were introduced on the basis of the $^{13}C\alpha$ and $^{13}C\beta$ chemical shifts [S. Spera and A. Bax, *J. Am. Chem. Soc.* 113, 5490 (1991)].
15. Slowly exchanging NH protons were identified by recording a series ^1H-^{15}N Overbodenhausen correlation spectra [G. Bodenhausen and D. J. Ruben, *Chem. Phys. Lett.* 69, 185 (1980); A. Bax, M. Ikura, L. E. Kay, D. A. Torchia, R. Tschudin, *J. Magn. Reson.* 86, 304 (1990)] over a period of ~24 hours starting within 5 min of dissolving an unexchanged sample of lyophilized protein in D_2O.
16. K. Wüthrich, *NMR of Proteins and Nucleic Acids* (Wiley, New York, 1986).
17. G. M. Clore and A. M. Gronenborn, *Science* 252, 1390 (1991).
18. The coordinates of the 22 SA structures and of the restrained minimized mean structure, $\overline{(SA)}r$, together with the experimental restraints, have been deposited in the Brookhaven Protein Data Bank.
19. M. Nilges, G. M. Clore, A. M. Gronenborn, *FEBS Lett.* 229, 317 (1988).
20. The hybrid distance geometry–SA protocol of (*19*) makes use of the program XPLOR [A. T. Brünger, G. M. Clore, A. M. Gronenborn, M. Karplus, *Proc. Natl. Acad. Sci. U.S.A.* 83, 3810 (1986); A. T. Brünger, *XPLOR Version 3 Manual* (Yale University, New Haven, 1992)] incorporating a distance geometry module [J. Kuszczewski, M. Nilges, A. T. Brünger, *J. Biomol. NMR* 2, 33 (1992)]. The protocol involves first calculating an initial set of substructures incorporating only about one third of the atoms by projection from *n*-dimensional distance space into cartesian coordinate space, followed by SA with all atoms. The target function that is minimized during SA (as well as in conventional Powell minimization) comprises only quadratic harmonic potential terms for covalent geometry (that is, bonds, angles, planes, and chirality), square-well quadratic potentials for the experimental distance and torsion-angle restraints (*11*), and a quartic van der Waals repulsion term for the nonbonded contacts (*19*). All peptide bonds were restrained to be trans. There were no hydrogen-bonding, electrostatic, or 6-12 Lennard-Jones empirical potential energy terms in the target function.
21. B. R. Brooks *et al.*, *J. Comput. Chem.* 4, 1987 (1983).
22. D. Eisenberg and A. D. McLachlan, *Nature* 319, 199 (1986); L. Chiche, L. M. Gregoret, F. E. Cohen, P. A. Kollman, *Proc. Natl. Acad. Sci. U.S.A.* 87, 3240 (1990).
23. In addition to deducing the approximate location of the main secondary structure elements from a qualitative analysis of the sequential and medium-range NOE data involving backbone protons, Redfield *et al.* (*10*) also presented six preliminary structures of IL-4 based on a very limited set of long-range NOE data. The precision of these structures, which had a backbone atomic rms difference of 3.2 Å (2.7 Å for the four helices), was only sufficient to suggest that the approximate topology probably comprised a left-handed four-helix bundle. Thus, the relative orientation of the four helices in their preliminary structures was very poorly defined and the packing of the four helices could not be defined at all. The difference in quality between these preliminary structures and those presented in this report does not therefore reflect a minor quantitative improvement in precision, but rather reflects a qualitative difference similar in nature, for example, to that between a 7 and 2.5 Å resolution x-ray structure.
24. S. S. Abdel-Meguid *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 84, 6434 (1987).
25. A. M. de Vos, M. Ultsch, A. A. Kossiakoff, *Science* 255, 306 (1992).
26. K. Diedrichs, S. Jacques, T. Boone, P. A. Karplus, *J. Mol. Biol.* 21, 55 (1991); K. Diedrichs, T. Boone, P. A. Karplus, *Science* 254, 1779 (1991).
27. D. Cosman *et al.*, *Trends Biochem. Sci.* 15, 265 (1990); K. I. Arai *et al.*, *Annu. Rev. Biochem.* 59, 7983 (1990); J. F. Bazan, *Proc. Natl. Acad. Sci. U.S.A.* 87, 6934 (1990); S. Davies *et al.*, *Science* 253, 59 (1991); D. P. Gearing *et al.*, *EMBO J.* 10, 2839 (1991).
28. C. Carr, S. Aykent, N. M. Kimack, A. D. Levine, *Biochemistry* 30, 1515 (1990).
29. B. M. Curtis *et al.*, *Proteins* 11, 111 (1991).
30. L. S. Park, D. Friend, K. Grabstein, D. L. Urdal, *Proc. Natl. Acad. Sci. U.S.A.* 84, 1669 (1987); L. S. Park, D. Friend, H. M. Sassenfeld, D. L. Urdal, *J. Exp. Med.* 168, 476 (1987).
31. K. Grabstein *et al.*, *J. Exp. Med.* 163, 1405 (1986); T. Yokota *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 83, 5894 (1986).
32. M. Carson, *J. Mol. Graphics* 5, 103 (1987).

33. Supported by the AIDS Directed Anti-Viral Program of the Office of the Director of the National Institutes of Health (G.M.C. and A.M.G.).

16 March 1992; accepted 23 April 1992

Rational Design of Potent Antagonists to the Human Growth Hormone Receptor

Germaine Fuh, Brian C. Cunningham, Rikiro Fukunaga, Shigekazu Nagata, David V. Goeddel, James A. Wells*

A hybrid receptor was constructed that contained the extracellular binding domain of the human growth hormone (hGH) receptor linked to the transmembrane and intracellular domains of the murine granulocyte colony-stimulating factor receptor. Addition of hGH to a myeloid leukemia cell line (FDC-P1) that expressed the hybrid receptor caused proliferation of these cells. The mechanism for signal transduction of the hybrid receptor required dimerization because monoclonal antibodies to the hGH receptor were agonists whereas their monovalent fragments were not. Receptor dimerization occurs sequentially—a receptor binds to site 1 on hGH, and then a second receptor molecule binds to site 2 on hGH. On the basis of this sequential mechanism, which may occur in many other cytokine receptors, inactive hGH analogs were designed that were potent antagonists to hGH-induced cell proliferation. Such antagonists could be useful for treating clinical conditions of hGH excess, such as acromegaly.

Knowledge of the molecular basis for hormone action is key to the rational design of hormone agonists and antagonists. High-resolution mutational analysis (*1*, *2*) and x-ray crystallographic studies (*3*) have defined two sites on hGH for binding two molecules of the extracellular domain of its receptor (hGHbp) (*4*). Dimerization of the hGHbp occurs sequentially, such that a hGHbp molecule binds to site 1 and then a second hGHbp molecule binds to both site 2 on hGH and a site on the first hGHbp (Fig. 1). A thorough examination of the biological importance of this model has been precluded because of the lack of an adequate cellular signaling assay for hGH. Here, we constructed a sensitive, cell-based assay for hGH, investigated the mechanism for signal transduction, and applied the assay for the design of antagonists to the hGH receptor.

The hGH receptor belongs to a large family of receptors of hematopoietic origin (*5*) that includes the interleukin-3 (IL-3) and granulocyte colony-stimulating factor (G-CSF) receptors. An IL-3–dependent myeloid leukemia cell line (FDC-P1) transfected with the full-length murine G-CSF (mG-CSF) receptor is stimulated to proliferate by G-CSF without IL-3 (*6*). We constructed a hybrid receptor that contained the hGHbp linked to a portion of the mG-CSF receptor containing the three extracellular fibronectin repeats and the transmembrane and intracellular domains (*7*). The fibronectin domains do not participate in the binding of G-CSF but are required for efficient expression of the mG-CSF receptor (*6*).

Competitive displacement of ^{125}I-labeled hGH from hybrid receptors on whole cells was used to establish the affinity for hGH and the approximate number of receptors per cell (*8*). In several independent

G. Fuh, B. C. Cunningham, J. A. Wells, Department of Protein Engineering, Genentech, Inc., 460 Point San Bruno Boulevard, South San Francisco, CA 94080.
R. Fukunaga and S. Nagata, Osaka Bioscience Institute, 6-2-4 Furuedai, Suita-shi, Osaka 565, Japan.
D. V. Goeddel, Department of Molecular Biology, Genentech, Inc., 460 Point San Bruno Boulevard, South San Francisco, CA 94080.
*To whom correspondence should be addressed.

binding experiments, the apparent dissociation constant (K_d) value for hGH was 0.1 ± 0.03 nM, and there were 1000 ± 300 receptors per cell. This affinity is about three to four times stronger than that for hGH binding to the soluble hGHbp and may reflect a high local concentration of receptors on cells (an avidity effect). Nontransfected cells lacked specific binding sites for hGH (9). At low concentrations, hGH induces cell proliferation with a median effective concentration (EC_{50}) of ~20 pM (Fig. 2A), a value somewhat lower than the apparent K_d for binding to whole cells (~100 pM). This may indicate that signaling for maximal cell proliferation requires less than total receptor occupancy.

Each hGH molecule is bivalent because it contains two separate sites for binding the hGHbp (Fig. 1). In contrast, the hGHbp is effectively univalent because each site uses virtually the same determinants to bind to either site 1 or site 2 on hGH (3). Excess hGH will dissociate the hGH·$(hGHbp)_2$ complex to form a hGH·hGHbp complex in which hGH is bound exclusively at site 1 to the hGHbp (1). Thus, excess hGH should antagonize signaling by preventing dimerization (Fig. 1). Indeed, at very high hGH concentrations the proliferation activity is lost [concentration required to inhibit proliferation by 50% (IC_{50}) ≅ 2 μM]. Cell proliferation induced by IL-3 was not altered in the presence of high concentrations of hGH (8 μM); thus, 8 μM hGH is not toxic to cells (9). This effect appears not to involve cross-linking of receptors between cells or other cell-to-cell interactions because the effects of hGH were not influenced by cell density. Furthermore, the assay is specific because FDC-P1 cells that contain the full-length mG-CSF receptor do not respond to hGH and cells that contain the hybrid receptor do not respond to G-CSF (10).

To further investigate the requirement

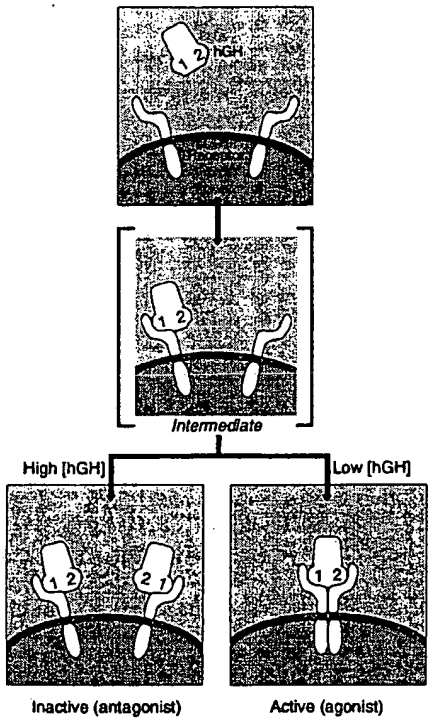


Fig. 1. Sequential dimerization model for activation of the hGH–mG-CSF hybrid receptor. At low concentrations, hGH binds first at site 1 and subsequently at site 2 (as indicated) to produce an active hGH·$(hGHbp)_2$ complex. At high concentrations, hGH saturates the receptor through site 1 interactions and acts as an antagonist. We show the receptors dissociated initially because, in the absence of hGH, the hGHbp does not self-dimerize as shown by ultracentrifugation for concentrations <0.1 mM. Nonetheless, it is possible that some full-length receptors are loosely pre-dimerized and become activated upon sequential binding of hGH.

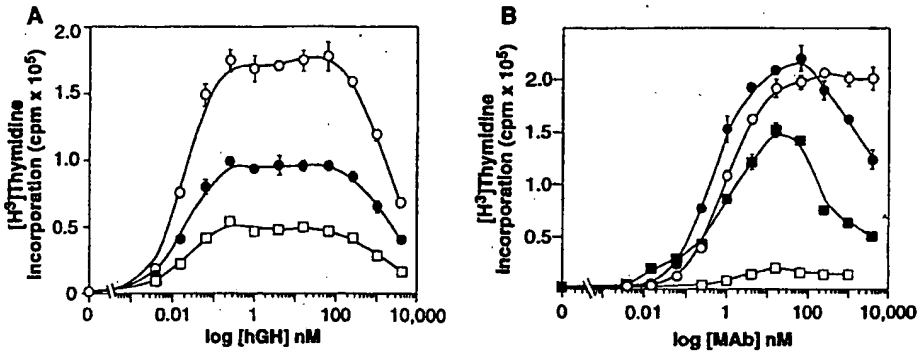


Fig. 2. Proliferation of FDC-P1 cells containing the hGH–mG-CSF hybrid receptor induced by hGH (7) (**A**) or induced by MAbs to the hGH receptor (**B**) (17). In (A), cells were grown in RPMI 1640 media supplemented with IL-3 (10 U/ml), 10 μM β-mercaptoethanol, and 10% FBS at 37°C and 5% CO_2 (6). Cells were washed with the same medium without IL-3. Cells were added to 96-well plates at a density of 4 × 10^5 cells per milliliter (○), 2 × 10^5 cells per milliliter (●), and 1 × 10^5 cells per milliliter (□) in 100 μl. Cells were then treated with various concentrations of hGH for 18 hours. To measure DNA synthesis, we added ^{3}H-labeled thymidine (1 μCi per well) to each well. After 4 hours, cells were collected and washed on glass filters. Scintillation cocktail (2 ml) was added, and radioactivity was counted with a Beckman LS1701 scintillation counter. In (B), cells were cultured as in (A) and plated at a density of 4 × 10^5 cells per milliliter in medium containing various concentrations of anti-hGH receptor MAb 263 (●), MAb 13E1 (○), MAb 3D9 (■), or MAb 5 (□). After 18 hours at 37°C, cells were washed, and the amount of [^{3}H]DNA synthesized was determined by scintillation counting as in (A). Each data point represents the mean of triplicate determinations, and error bars indicate the SD.

Table 1. Summary of dose-response data for a variety of anti-hGH receptor MAbs, FAbs (17), and hGH mutants (12) for stimulating proliferation of FDC-P1 cells containing the hGH–mG-CSF hybrid receptor. "None" indicates that no effect was observed; ND, not determined. K_d values for MAbs binding to the hGHbp were taken from (14). The K_d values for hGH and variants were measured with a competitive displacement assay in which ^{125}I-labeled hormone bound to hGHbp was precipitated with MAb 5 (2, 13). This gives the affinity for the monomeric hGH·hGHbp complex. Values for EC_{50} were taken from titration curves shown in Fig. 2, A and B, and Fig. 4 and represent the half-maximal concentration for stimulation of cell proliferation. Data are the mean of triplicate assays, and the SDs were within 15% of mean. Values shown with > indicate that maximal stimulation or inhibition of proliferation was not detected at the concentrations tested. For these cases, we report only estimates of the EC_{50}. IC_{50} refers to the concentration leading to 50% inhibition of maximal cell proliferation.

Protein	K_d (nM)	EC_{50}	IC_{50} (self-antagonism)
MAb 263	0.6	0.3 nM	~3 μM
MAb 13E1	3.2	0.8 nM	>10 μM
MAb 3D9	2.2	0.8 nM	0.2 μM
MAb 5	0.7	~2.5 nM	>1 μM
FAb 263	ND	>1.5 μM	ND
FAb 13E1	ND	>3 μM	ND
FAb 3D9	ND	>0.1 μM	ND
FAb 5	ND	>1 μM	ND
hGH	0.3	20 pM	2 μM
K172A/F176A	200	25 nM	None
G120R	0.3	None	None
H21A/R64K/E174A	0.01	20 pM	60 nM
H21A/R64K/E174A/G120R	0.01	None	None

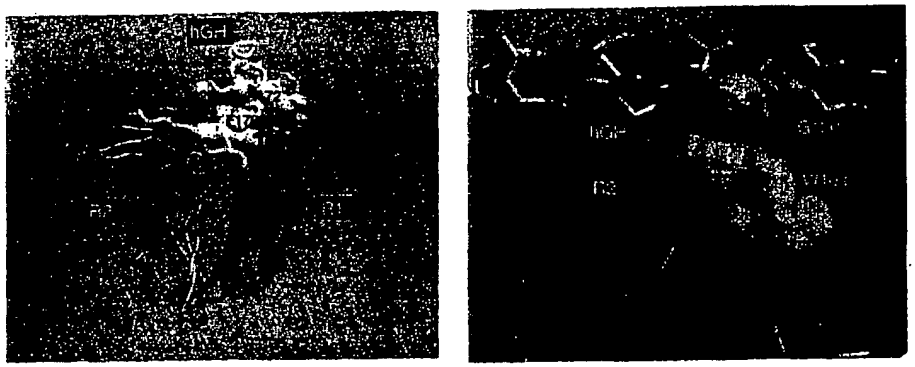


Fig. 3. Molecular models based on x-ray crystallography of hGH bound to the hGHbp (*3*). (**A**) A ribbon diagram of hGH (white) bound to two molecules of the extracellular domain of the hGH receptor (hGHbp; gray and black). The α carbon positions of mutant residues in hGH are shown by black dots. K172 and F176 are located in site 1 (interface with black hGHbp) and G120 is located in site 2 (interface with gray hGHbp). Gray dots indicate structures in the hGHbp that are not well defined by the electron density. The model is based on a 2.7 Å resolution x-ray structure of the complex (*3*). (**B**) A close-up showing that G120 located on helix 3 of hGH makes van der Waals contact with W104 from the hGHbp bound to site 2. R1 and R2, receptor 1 and receptor 2, respectively.

for dimerization of the hGHbp to signal in the hybrid receptor cell proliferation assay, we used bivalent monoclonal antibodies (MAbs) and univalent fragments derived from them (FAbs) that recognized the hGHbp. At low concentrations, three of four different MAbs to the receptor were as potent as hGH in inducing cell proliferation (Fig. 2B and Table 1). The EC_{50} value for each MAb (0.3 to 1 nM) was usually somewhat less than the K_d value determined by enzyme-linked immunosorbent assay (Table 1). As with hGH, this may reflect avidity effects on whole cells or that maximal signaling is achieved at less than 100% receptor occupancy, or both. At much higher concentrations (0.2 to ~3 μM), two of these MAbs were less effective at stimulating proliferation, presumably because excess MAb blocks receptor cross-linking by binding monovalently to hGHbp. Corresponding monovalent FAb fragments had little or no effect on cell proliferation (Table 1), which indicates further that bivalency is required for signaling activity.

The differences in stimulation of cell proliferation at low concentrations and inhibition at high concentrations for these MAbs (Fig. 2B) can be explained by the different ways they bind to the hGHbp. MAb 5 prevents binding of a second hGHbp to the hGH·hGHbp complex (*1*), possibly by binding to the region where both receptors contact each other (Fig. 1). The fact that MAb 5 is the least efficient at stimulating proliferation may indicate that the receptors need to approach each other closely for optimal signaling. MAb 13E1 did not inhibit proliferation at the concentrations tested. This MAb blocks hGH binding (*11*) and probably binds like hGH to form very stable receptor dimers. In contrast, MAbs 263 and 3D9 bind at sites away from the hormone-receptor interfaces (*11*) and show similar agonistic and antagonistic effects on proliferation. Maximal stimulation of proliferation by hGH occurred over a wider range of concentrations than did maximal stimulation by MAbs 263 and 3D9, perhaps because with hGH bound, the dimers have the optimal receptor-to-receptor contacts. The fact that MAbs 263 and 3D9 are agonists suggests that the structural constraints for formation of active dimers are rather loose.

FAb fragments derived from MAb 13E1 or MAb 5 antagonized hGH-induced cell proliferation, whereas those derived from MAbs 263 and 3D9 did not (Table 2). These studies are consistent with the fact that the binding of MAb 13E1 or MAb 5 to their epitopes blocks hormone-to-receptor or receptor-to-receptor interfaces, respectively.

Table 2. Summary of antagonist effects of FAbs and hGH analogs that block hGH-induced cell proliferation of FDC-P1 cells containing the hybrid hGH–mG-CSF receptor. Cells were incubated with 1 nM hGH and various concentrations of FAb (*17*) or hGH analog (*12*). The IC_{50} is the concentration required to block 50% of the cell proliferation activity of hGH. "None" indicates that no inhibition was observed at concentrations of FAb or hGH analog of up to 10 μM.

Protein	IC_{50}
FAb 263	None
FAb 13E1	0.8 μM
FAb 5	0.2 μM
FAb 3D9	None
hGH	2 μM
K172A/F176A	None
G120R	20 nM
H21A/R64K/E174A	60 nM
H21A/R64K/E174A/G120R	2 nM

To determine the structural requirements for dimerization of hGH (Fig. 1), we examined mutants of hGH that were designed to reduce binding of the receptor to site 1 or site 2 (Fig. 3). The mutant K172A/F176A (*12*), which preserves site 2 determinants but alters important side chains in site 1, promoted cell proliferation, but the EC_{50} was shifted to a concentration about 10^3 times higher than that of wild-type hGH (Fig. 4 and Table 1). This is consistent with the 560-fold reduction in the affinity for site 1 binding of the K172A/F176A mutant as compared to that of the wild-type hGH when measured in vitro (*13*). No inhibition of proliferation with K172A/F176A was observed at the concentrations tested.

On the basis of the x-ray structure of the hGH·$(hGHbp)_2$ complex (*3*), we designed

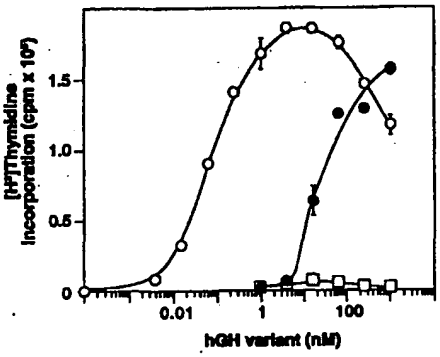


Fig. 4. Proliferation of FDC-P1 cells containing the hGH–mG-CSF hybrid receptor, caused by increasing concentrations of wild-type hGH (○), the site 1 hGH variant K172A/F176A (●), and the site 2 hGH variant G120R (□). Cells were cultured, treated, and assayed as described in Fig. 2A, except that cells were treated for 18 hours with [³H]thymidine. The hGH mutants were prepared and purified as described (*2*, *12*).

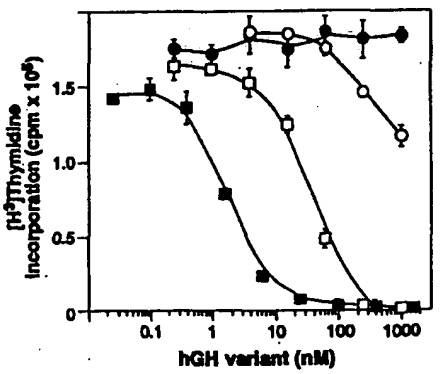


Fig. 5. Antagonism of hGH-induced cell proliferation by hGH variants. Cells were prepared as in Fig. 2A and incubated with 1 nM hGH and various concentrations of the site 1 mutant K172A/F16A (●), the site 2 mutant G120R (□), the combined enhanced site 1 and site 2 mutant (H21A/R64K/E174A/G120R) (■), and wild-type hGH (○).

a mutant G120R, which retains a functional site 1 but on which site 2 is sterically blocked (Fig. 3B). This variant did not affect cell proliferation at the concentrations tested (Fig. 4). Thus, binding to either site 1 or site 2 is necessary but not sufficient for promoting cell proliferation.

If the sequential signaling mechanism (Fig. 1) is correct, mutants blocked in site 2 binding (but not in site 1 binding) should antagonize hGH-induced cell proliferation. To test this, we cultured cells with enough hGH (1 nM) to support 90% of maximal cell proliferation and added increasing concentrations of wild-type hGH or the mutants in site 1 (K172A/F176A) or site 2 (G120R). As expected, the site 2 mutant antagonized hGH whereas the site 1 mutant was ineffective (Fig. 5). In fact, the site 2 mutant was nearly 100 times more potent as an antagonist (IC_{50} = 20 nM) (Table 2) than wild-type hGH (IC_{50} = 2 μM). For hGH to be antagonistic, free hormone must react with free receptors before the hGH-bound intermediate does so. This only occurs at high concentrations of hGH. In contrast, once G120R is bound, it cannot dimerize and agonize the receptor. Thus, G120R as an antagonist does not need to compete against G120R as an agonist.

Although G120R is a much more potent antagonist than hGH, 20 nM G120R was required to inhibit by 50% the proliferative effect of 1 nM hGH (Table 2). This may reflect the fact that hGH is bound through interaction of sites 1 and 2 with two receptors more tightly than G120R is bound in the complex with a single receptor through site 1 alone. Furthermore, maximal signaling by hGH may not require 100% receptor occupancy. In either case, improving the affinity of site 1 for hGHbp in the G120R mutant should make it a more potent antagonist.

Single-site hGH variants have been produced (*2*, *14*) that bind more tightly to the hGHbp at site 1. A variant that contains all three of these mutations (H21A/R64K/E174A) bound 30 times more tightly than wild-type hGH to the hGHbp (Table 1). This variant had an IC_{50} for inhibiting proliferation that was about 30 times lower than that of hGH. This is consistent with the notion that the inhibitory effect results from competition for binding to hGHbp between site 2 on the bound hormone-receptor intermediate and the free site 1 on the soluble hormone. The fact that improvement in site 1 binding affinity did not improve the efficacy of the hormone as an agonist may be understood upon future analysis of the on and off rates.

We further mutated this variant by changing Gly^{120} to Arg. The mutant with all four modifications was ten times more potent than G120R as an hGH antagonist (Fig. 5 and Table 2). This is further evidence for the importance of site 1 binding properties for antagonism.

Our data suggest that the inhibition of proliferation caused by hGH, MAbs, and their derivatives is the result of blocking receptor dimerization rather than causing down-regulation of receptors. First, cells propagated with IL-3 instead of hGH do not show a greater hGH response or hGH receptor number (*9*). Second, receptor down-regulation is usually correlated to receptor activation. The ratio of EC_{50} to IC_{50} for each of the MAbs and hGH varies widely, which shows that receptor activation can be readily uncoupled from inhibition by the alteration of binding properties. Finally, the G120R mutant is inactive as an agonist, although it is a more potent antagonist than hGH (Fig. 5), and pretreatment of cells with G120R does not enhance its antagonistic effect (*9*). Thus, the antagonistic effect of G120R is not consistent with receptor down-regulation. It is possible that the inhibitory effects observed for other hormones at high concentrations may occur because receptor dimerization is blocked by self-competition.

Our studies indicate that sequential dimerization is crucial for hybrid-receptor activation. Knowledge of this mechanism and the structural (*3*) and functional (*1*, *2*) properties of the binding interfaces allowed us to design potent antagonists to the hGH receptors, which may be useful in the clinical treatment of hGH excess acromegaly (*15*). In fact, a transgenic strain of mice that expresses large amounts of bovine GH altered in site 2 produces dwarf mice (*16*). This mechanism-based strategy for design of potent antagonists for hGH may be applicable to other hormones such as prolactin, placental lactogen, IL-2, IL-3, IL-6, G-CSF, granulocyte-macrophage–CSF, erythropoietin, and related hematopoietines and cytokines (*5*) if sequential binding of two receptors to a single hormone molecule is required for their signaling.

REFERENCES AND NOTES

1. B. C. Cunningham *et al.*, *Science* 254, 821 (1991).
2. B. C. Cunningham and J. A. Wells, *ibid.* 244, 1081 (1989).
3. A. M. de Vos, M. Ultsch, A. A. Kossiakoff, *ibid.* 255, 306 (1992).
4. J. A. Wells *et al.*, *Recent Prog. Horm. Res.*, in press.
5. J. F. Bazan, *Proc. Natl. Acad. Sci. U.S.A.* 87, 6934 (1990); D. Cosman *et al.*, *Trends Biochem. Sci.* 15, 265 (1990); L. Patthy, *Cell* 61, 13 (1990).
6. R. Fukunaga, E. Ishizaka-Ikeda, S. Nagata, *J. Biol. Chem.* 265, 14008 (1990); R. Fukunaga, E. Ishizaka-Ikeda, C.-X. Pan, Y. Seto, S. Nagata, *EMBO J.* 10, 2855 (1991).
7. A hybrid receptor was constructed from cDNA that contained exons 1 through 7 of the hGH receptor (which encode the secretory signal and the extracellular hGH binding domains) linked to exons 9 through 17 of the mG-CSF receptor (which encode the three fibronectin domains and the entire transmembrane and intracellular domains). The protein encoded by the hybrid cDNA contains the amino acids −18 to 243 of the hGH receptor and 309 to 812 of the mG-CSF receptor. In this construction, Cys^{241} was replaced by Arg, which has no effect on binding to hGH [G. Fuh *et al.*, *J. Biol. Chem.* 265, 3111 (1990)]. Sequences derived from the hGH receptor [D. W. Leung *et al.*, *Nature* 330, 537 (1987)] and from the mG-CSF receptor [R. Fukunaga *et al.*, *Cell* 61, 341 (1990)] were cloned by means of the polymerase chain reaction [R. Higuchi, in *PCR Protocols: A Guide to Methods and Applications*, M. Innis *et al.*, Eds. (Academic Press, New York, 1989), pp. 177–183] into the vector pER-BOS (*6*). The hybrid cDNA was introduced into mouse FDC-P1 cells as described (*6*). Stable transformants that expressed the hybrid protein were identified by their ability to bind ^{125}I-hGH (*8*).
8. Cells grown with IL-3 were washed before the assay with phosphate-buffered saline (PBS) that contained 10% fetal bovine serum. Cells (1.2 × 10^6 per milliliter) were incubated with serial dilutions of hGH in the presence of ^{125}I-labeled hGH variant Y103A (20 pM) for 18 hours at 4°C. The Y103A variant was used to prevent iodination of Y103, which inhibits the binding of the receptor to site 2 (*9*). Cells were washed with PBS twice to remove the excess hormone, and the bound radioactivity was counted.
9. G. Fuh and J. A. Wells, unpublished results.
10. E. Ishizaka-Ikeda, R. Fukunaga, S. Nagata, unpublished results.
11. S. H. Bass, M. G. Mulkerrin, J. A. Wells, *Proc. Natl. Acad. Sci. U.S.A.* 88, 4498 (1991).
12. Mutants of hGH were prepared by site-directed mutagenesis [T. A. Kunkel, J. D. Roberts, R. A. Zakour, *Methods Enzymol.* 154, 367 (1987)] and purified as described (*2*, *13*). Mutants are designated by the wild-type residue followed by its position and the mutant residue in the single-letter code A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Try; and Y, Tyr. Thus, G120R indicates that Gly^{120} was converted to Arg. Multiple mutants are represented by a series of single mutants separated by slashes.
13. B. C. Cunningham and J. A. Wells, *Proc. Natl. Acad. Sci. U.S.A.* 88, 3407 (1991).
14. B. C. Cunningham, D. J. Henner, J. A. Wells, *Science* 247, 1461 (1990).
15. S. Melmed, *N. Engl. J. Med.* 322, 966 (1990); L. A. Frohman, *J. Clin. Endocrinol. Metab.* 72, 1175 (1991).
16. W. Y. Chen, D. C. Wright, B. V. Mehta, T. E. Wagner, J. J. Kopchick, *Mol. Endocrinol.* 5, 1845 (1991).
17. The MAbs 5 and 263 were from Agen (Parsippany, NJ) and have been described [R. Barnard, P. G. Bundesen, D. B. Rylatt, M. J. Waters, *Endocrinology* 115, 1805 (1984); *Biochem. J.* 231, 459 (1985)]. The MAbs 13E1 and 3D9 were from the Genentech hybridoma group, and their properties have been described (*1*, *11*). Briefly, MAbs were purified from mouse ascites fluid by binding to Protein A–Sepharose (Pharmacia) and eluted with 0.1 M acetate (pH 3.0). FAb fragments were prepared by treating MAbs with dithiothreitol-activated papain (1 part to 50 parts MAb, by weight) in PBS with 10 mM cysteine for 1 hour. Digestions were stopped by adding 0.1 M iodoacetamide. The Fc and residual MAb were removed by adsorption onto Protein A–Sepharose twice, and the FAb fragments were further purified by gel filtration on Superose 12 (Pharmacia).

18. We thank B. de Vos for providing x-ray coordinates for the hGH·$(hGHbp)_2$ complex and for useful comments on the manuscript, E. Ishizaka-Ikeda for help in constructing the hybrid hGH–G-CSF receptor, T. Hynes for help with molecular modeling, the oligonucleotide synthesis group at Genentech for DNA synthesis, and M. Thorner and C. Silva for sharing data on hGH stimulation of IM-9 cells before publication.

6 March 1992; accepted 1 May 1992

DAVID J AUSTIN, GERALD R CRABTREE AND STUART L SCHREIBER

Proximity versus allostery: the role of regulated protein dimerization in biology

Regulated dimerization of proteins is increasingly understood to be important in many cellular processes, including signaling, transcription and protein degradation. Organic molecules that induce dimerization may offer as much potential to regulate biological processes as those that allosterically induce conformational change.

Chemistry & Biology November 1994, 1:131–136

Several different mechanisms are used to transfer information in biological systems. The most familiar one is that of induced conformational change, relying on the ability of ligands to induce an allosteric change in their receptors. But a second mechanism for information transfer, only relatively recently uncovered by cell biologists, is equally important. This is the regulated association of specific proteins, 'protein dimerization'. A new class of organic molecules that can induce the association of specific proteins can be envisaged; such molecules may allow the regulation of biological systems, like the classical allosteric agents that have been a primary focus of biological and medical research for many years.

Cell-surface receptors

The two processes of allosteric change and receptor dimerization are clearly contrasted in two types of cell surface receptors that activate intracellular signaling pathways, the G-protein coupled receptors and the growth factor or growth hormone receptors (Fig. 1). Both activate signaling pathways in response to external binding events, using distinctly different mechanisms. The G-protein coupled receptors are allosterically activated when ligands bind to their transmembrane domain. This ligand-induced conformational change allows the cytoplasmic loops to activate an associated GTP-binding protein (G-protein). Receptor-associated G-proteins are trimeric, consisting of an α subunit that binds guanine nucleotides and β and γ subunits; the α and $\beta\gamma$ subunits have independent signaling functions. The ligand-induced change in receptor conformation is thought to promote the release of GDP so that GTP can bind in its place. Once GTP is bound, the activated Gα-GTP complex and the $\beta\gamma$ subunits separate from each other and dissociate from the membrane, then propagate the signal further by interacting with downstream effectors [1].

The growth-factor-type receptors are activated by a ligand-induced protein dimerization. Hormones and cytokines such as erythropoietin, granulocyte colony-stimulating factor and human growth hormone (hGH) can cross-link two receptors, resulting in the juxtaposition of two cytoplasmic tails. Many of the dimerization-activated receptors (for example, the receptor tyrosine kinases) have protein kinase domains within their cytoplasmic tails that phosphorylate the neighboring tail upon dimerization. In related receptors, such as the cytokine receptors, the cytoplasmic tails lack intrinsic kinase activity but are functionally similar since they associate with protein kinases. In both cases, the phosphorylation results in the activation of a signaling pathway (see below).

In the case of hGH and its receptor, structural studies have revealed the basis for the dimerization of the extracellular domain, which of course causes the dimerization of the intracellular domain. For this receptor, dimerization brings together the associated JAK protein tyrosine kinases. hGH

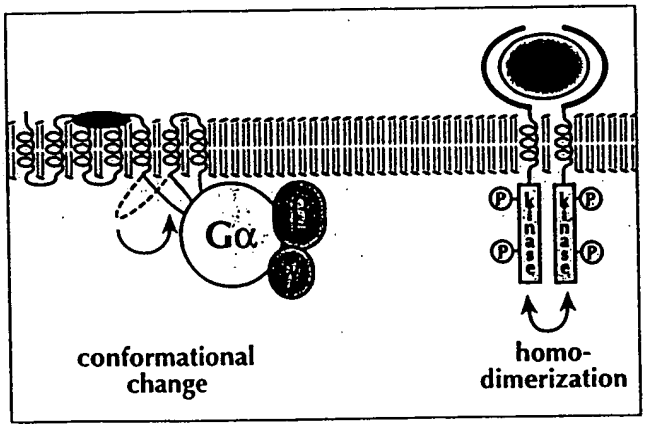


Fig. 1. Allostery and proximity. The allosteric effect of a low molecular weight ligand such as a neurotransmitter on its G-protein-coupled (serpentine or seven-transmembrane spanning) receptor and the dimerizing effect of an extracellular protein such as a growth factor, cytokine or hormone on its normally monomeric receptor. When a G-protein-coupled receptor binds its ligand, a conformational change in the intracellular cytoplasmic loops of the receptor results in the activation of a trimeric G-protein, resulting in GDP/GTP exchange and the dissociation of the Gα subunit/GTP complex from the $\beta\gamma$ subunits. When a growth factor binds to its receptor, however, two receptors are crosslinked. If the cytoplasmic domain of the receptor contains a kinase domain, the two receptor tails will phosphorylate each other.

binds to the extracellular portion of the hGH-receptor with 1:2 stoichiometry by a homodimerization mechanism [2]. The determination of the structure of hGH bound to two extracellular domains provides a fascinating insight into the workings of a natural homodimerizer (Fig. 2) [3]. Although the two receptor molecules that are brought together by a single molecule of hGH are identical, there is no symmetry in the hGH molecule. The two binding events can be clearly distinguished, and must occur in the correct order. The first binding event, although it does not change the shape of the hGH molecule, creates a composite surface that can now bind to the second hGH receptor. Thus, binding at site 1 is essential to allow binding at site 2. By analogy with hGH, we define a dimerizer as an agent that has its effect by bringing two molecules together by binding to both of them. Although it may form only part of the binding site for the second molecule, it does not, or at least need not, change the shape of either molecule to have its effect.

Allosteric conformational change is important in cell-membrane signaling pathways, the control of transcription, and the activity of several enzymes, to name just a few examples. But regulated protein dimerization is important in at least as many different classes of cellular processes, including diverse signaling pathways, immunological recognition, transcription and the control of protein degradation. We will first examine a few examples of the importance of allosteric change before turning to areas in which regulated dimerization is important.

Allostery

An allosteric conformational change is one that is initiated by the binding of a ligand (even a small molecule) to a protein or macromolecule, but takes place at a location distant from the binding site. In most cases this ligand-induced conformational change influences the activity or function of the protein or macromolecule. Nature and synthetic chemists have both provided many low molecular-weight molecules that can induce allosteric changes in their receptors with dramatic cellular consequences.

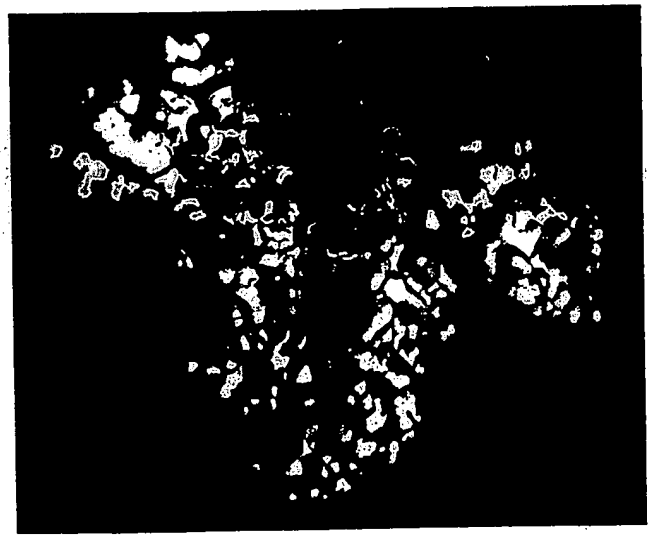


Fig. 2. The homodimerizer human growth hormone (hGH, red) bound to two identical extracellular domains of its receptor (blue). The receptor domains are shown as molecular surfaces, and a sampling of the hGH protein side chains is also shown. A remarkable feature of the hGH homodimerizer is that it has two distinct receptor-binding surfaces that each bind the hGH receptor differently, yet the binding site used by the receptor is essentially the same in both cases. Generated using the program GRASP [20].

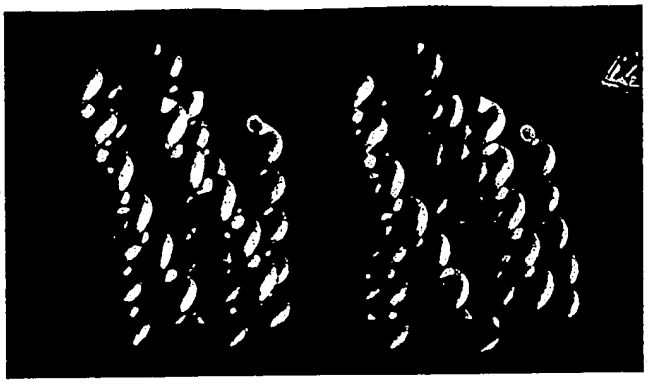


Fig. 3. Bacteriorhodopsin's seven-helical bundle, allosterically regulated by all-*trans*-retinal (left) and 13-*cis*-retinal (right). These are theoretical models based on atomic coordinates determined by electron diffraction and subsequent computational energy optimizations [4]. It is believed that the seven-transmembrane receptors function in similar ways. Generated using GRASP [20].

Retinal is an example of a molecular switch acting on the halobacterial proton pump, bacteriorhodopsin. Bacteriorhodopsin consists of a seven-helix transmembrane protein, but is not a receptor and is not linked to a trimeric G-protein. Instead, it is covalently bound to all-*trans*-retinal, which upon photolysis isomerizes to 13-*cis*-retinal. This event results in a conformational change in the protein's helical bundle, opening an ion channel and inducing a proton concentration gradient across the cellular membrane (Fig. 3; [4]). This gradient drives ATP synthesis in the bacterium.

An elegant example of allosteric protein activation is the metabolic self-regulation of tryptophan biosynthesis. The *trp* repressor is a dimeric protein that regulates the transcription of the *trp* operon, which encodes the tryptophan biosynthesis genes. In the absence of tryptophan, the protein is inactive and the *trp* operon genes are transcribed, allowing the biosynthesis of tryptophan. Once the protein binds to tryptophan, it changes its conformation, and becomes activated (Fig. 4; [5,6]). The active conformation of the repressor allows it to bind specifically to the operator region of the *trp* operon. When the operator site is bound, the promoter region is blocked, preventing transcription. Therefore, in the presence of excess tryptophan, tryptophan biosynthesis is turned off.

The anti-HIV agent nevaripine [7] is a synthetic molecule that inhibits reverse transcription of viral RNA mediated by HIV-1 reverse transcriptase (RT). The proposed mode of action is based partly on the crystal

structure of Steitz and coworkers (Fig. 5; [8]), which shows that the RNA template that RT copies into DNA binds in a large cleft with a shape resembling that of a right hand. The subdomains of the cleft are termed the 'fingers', 'palm' and 'thumb', and the function of the enzyme appears to involve a conformational change of the thumb with respect to the palm and fingers. The nevaripine-bound conformation of HIV-1 reverse transcriptase is thought to inhibit polymerase activity by crimping or restricting the motion of the 'thumb'.

Thus allosteric change occurs in diverse types of proteins, is induced by diverse ligands, and has a variety of effects. Such allosteric changes have often been the target of drug development efforts, and agents that affect them will no doubt continue to be important in improving our understanding of cell biology as well as in medicine.

Fig. 5. Nevaripine (yellow) bound to the p66 domain (blue) of HIV-1 reverse transcriptase. The p51 domain (red) serves as a scaffolding for the p66 domain and the RNase H domain (white). The orange spheres represent the polymerase active site and the lavender balls show the RNase H active site. Modeled into the structure is single-stranded RNA (green) and single-stranded DNA (white). Courtesy of Dr Julian Adams, Myogenics Inc., Cambridge, MA. The action of the 'thumb' (arrow) is thought to be essential for DNA polymerization activity, and to be blocked by nevaripine.

Induced proximity

The homodimerizing action of growth factors and hormones, of which hGH is one example, sets in motion intracellular signaling pathways that also depend on the proximity induced by regulated protein dimerization. Many growth-factor receptors activate the well-studied Ras pathway, in which several contingent protein heterodimerizations result in the translocation of signaling proteins to the inner leaflet of the plasma membrane (Fig. 6; [9]). The first step in the pathway initiated by homodimerization of the epidermal growth factor (EGF) receptor, for example, is *trans*-phosphorylation of tyrosines in the tail of the membrane receptor. The phosphorylated receptor tail can now bind the protein Grb2, which contains one SH2 domain that binds phosphotyrosine residues in the receptor tail [9]. The molecular details of this association are now understood (Fig. 7; [10]). Grb2 is a heterodimerizing agent; as well as the SH2 domain that permits association with phosphorylated receptor tails, it also contains two SH3 domains that bind a number of proteins including an activator of Ras named Sos. Ras is a GTP-binding protein, but is not one of the class of heterotrimeric G proteins that is directly activated by the seven-transmembrane spanning receptors shown in Fig. 1; however, it is activated in the same way, by the exchange of GDP for GTP. Ras carries farnesyl and palmitoyl groups, and therefore resides in the inner membrane; when Grb2 moves to the membrane with Sos, it thus brings Sos into proximity with Ras. Sos can then promote the exchange of GTP for GDP, activating Ras. Unlike the G-protein coupled receptors, which require allosteric activation to promote GDP release from the heterotrimeric G proteins, Sos only needs to be brought close to its target to initiate signalling. Thus all that is necessary for signaling is recruitment of Sos to the membrane by a heterodimerizing agent, in this case Grb2 [11].

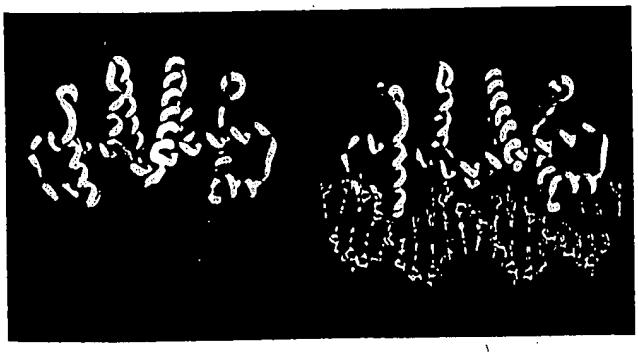


Fig. 4. Turning off transcription. The *trp* aporepressor (left) and the altered conformation of the activated repressor (right) bound to tryptophan (green) and duplex DNA. Generated using GRASP [20].

Organic heterodimerizers

Once Ras is bound to GTP, yet another inducible protein-dimerization event results. Activated Ras binds to the serine/threonine kinase Raf and recruits it to the inner membrane; at present, it is unclear how GTP binding allows Ras to bind to Raf. One possibility is that GTP acts like a heterodimerizer, binding simultaneously to both Ras and Raf.

An analogous situation exists with the low-molecular-weight, organic heterodimerizers cyclosporin A (CsA) and FK506 (Fig. 8) [12]. By binding tightly to the soluble, cytosolic protein FKBP and forming a composite surface, FK506 recruits it to the protein phosphatase calcineurin, which is associated with intracellular membrane components via its myristic acid moiety. CsA similarly binds cyclophilin and calcineurin simultaneously, and the result in both cases is the inhibition of calcineurin function, which is required for intracellular transmission of the signal that emanates from the activated T cell receptor. Because of the immunosuppressive effects of their ligands, FKBP and cyclophilin are known as immunophilins.

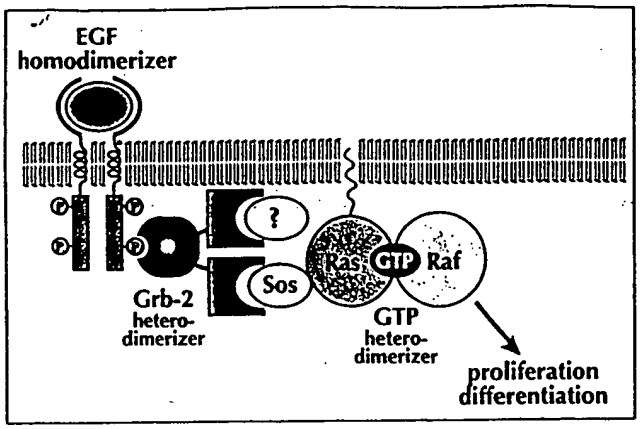


Fig. 6. The early steps of the Ras pathway, illustrating the role of homodimerizers and heterodimerizers. Once the cytoplasmic domain of the growth factor receptor has been phosphorylated (see Fig. 1), it can bind the heterodimerizer Grb-2. This in turn binds Sos, a guanine nucleotide exchange factor for Ras, bringing it to the membrane and into proximity with Ras, which is held near the membrane by farnesyl and palmitoyl groups. The exchange of GDP for GTP creates a binding site for Raf; thus, GTP in this context can be considered to be a heterodimerizing agent.

CsA and FK506 clearly make contact both with the immunophilin and with calcineurin. Like the hGH receptor homodimerization induced by hGH, however, heterodimerizer binding to the second protein is dependent on the first binding event. Thus, CsA does not bind to calcineurin unless it has already associated with cyclophilin, nor can FK506 bind to calcineurin in the absence of FKBP. In both cases, a composite surface composed of both the immunophilin and its organic ligand is formed that is recognized by calcineurin. These natural products and others that, like them, can have a biological effect by bringing two proteins into close proximity, fall into a new class of 'chemical inducers of dimerization' (CID) that seems certain to grow in the future.

The view of these natural products as small molecule heterodimerizers, equipped with two protein-binding surfaces, reveals another fascinating aspect of natural dimerizing agents. They can cause a single protein to bind to multiple protein partners, with specificity determined by the dimerizer. An example is seen in the case of FKBP12. This immunophilin binds to another natural product and heterodimerizer, rapamycin. Rapamycin binding creates a composite surface distinct from the one formed with FK506, resulting in the heterodimerization of another signaling protein, the putative lipid kinase FKBP–rapamycin-associated protein (FRAP) [13,14]. FKBP12 binds neither calcineurin nor FRAP, but in the presence of the heterodimerizers FK506 or rapamycin it can form two distinct receptor–ligand–receptor complexes (Fig. 9), FKBP12–FK506–calcineurin and FKBP12–rapamycin–FRAP. The former complex prevents resting cells from entering into the cell cycle, while the latter prevents cells from progressing through the first gap phase (G1) of the cell cycle.

Peptide dimerizers

The most remarkable illustration of the ability of dimerizers to increase the number of targets an individual protein can bind to is surely that of the MHC class I and II antigen-presenting molecules. Any given member of this family can bind many different peptides; these peptides are routinely produced by degradation of proteins within the cell. In most cases the peptide is derived from a self protein, but in some cases it results from the degradation of a viral or bacterial peptide. MHC class I and II molecules select peptides from this mixture of self and foreign peptides according to the preferences of their binding sites, and present them to T cells for identification. Depending on which peptide has been selected, the surface of the MHC-peptide complex will vary slightly in shape, which in turn will determine which of the many T cell receptors available will bind to the complex [15]. The antigenic peptides can be viewed as heterodimerizers that elicit a biological response by forming a composite surface with their MHC receptor, thereby inducing a specific protein–protein interaction that otherwise would not occur.

The importance of dimerizers and inducible proximity is not restricted to the early and intermediate events of signaling pathways. The culmination of many of these pathways is the activation of transcription in the nucleus, and here, again, regulated heterodimerization is important. Indeed, transcriptional activators themselves can be viewed as heterodimerizers composed of DNA-binding and protein-binding surfaces (Fig. 10). Class II nuclear genes have an element within their promoter region known as the TATA box, a short sequence composed entirely of T-A base-pairs. The TATA-box binding protein (TBP) is required for initiation of transcription of all such genes. TBP alone does not efficiently activate transcription, however; it requires a number of transcription factors and TBP-associated factors (TAFs).

Fig. 7. Crystal structure of the SH2 domain found in the non-receptor tyrosine kinase Src, bound to the phosphorylated peptide ligand pYEEI. The heterodimerizer Grb-2 is expected to bind its phosphorylated substrate in a similar way. This structure thus gives insight into one of the two protein–ligand interactions required for a natural dimerizer to function. Generated using GRASP [20].

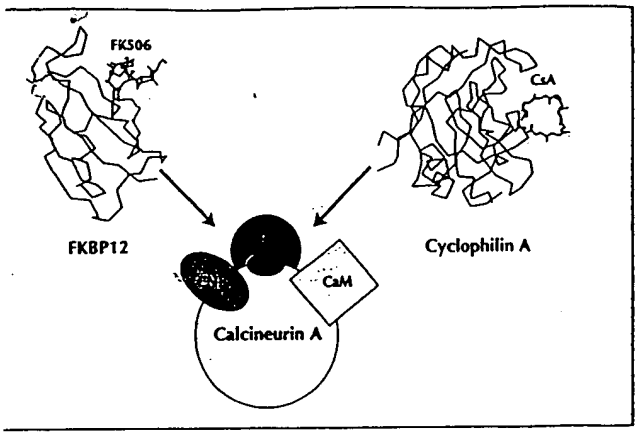


Fig. 8. FK506 and CsA both function as heterodimerizers by causing immunophilins to interact with the signaling protein calcineurin. These ligands create a composite surface following binding to their immunophilin receptors that results in the formation of high-affinity ligands to calcineurin. CNB, calcineurin subunit B; CaM, calmodulin; Imm, immunophilin; L, ligand.

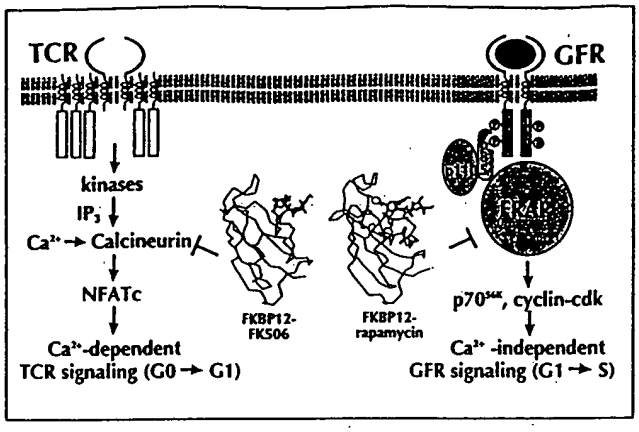


Fig. 9. FK506 and rapamycin both bind FKBP12, but each forms a different composite surface resulting in the heterodimerization of the immunophilin with different signaling proteins thereby resulting in different biological outcomes. In the T cell receptor (TCR) signaling pathway, engagement of the TCR by an appropriate MHC–peptide complex initiates signaling via inositol triphosphate (IP_3) and calcineurin, which eventually leads to the activation of the cytoplasmic component of the transcription factor NFAT ($NFAT_c$), and entry of cells into the cell cycle. The FKBP12–rapamycin complex, on the other hand, blocks the signaling pathway initiated by the IL-2 receptor and several other cytokine and growth factor receptors (GFRs) by binding to FKBP–rapamycin-assocated protein (FRAP). The FRAP signaling pathway activates the protein $p70^{S6K}$ and cyclin-dependent kinases (cdks), and causes the cell to make the transition from the first gap phase of the cell cycle (G1) to the synthesis (S) phase.

Enhancers that specifically regulate the transcription of these genes do so by binding to an enhancer sequence some distance from the site of initiation and recruiting a TBP-associated factor (TAF) to the vicinity of the TATA box, facilitating the complete formation of an initiation complex. Thus, the activator uses its ability to bind both DNA and TAF to bring the latter into close proximity with the neighboring promoter sequence [16].

Although signaling pathways are perhaps the richest source of examples of the induction of proximity using molecular dimerizers, this mechanism of information transfer is not limited to signaling. We have already seen that proteins can be caused to translocate to various parts of the cell by dimerizers. The recruitment of signaling molecules to the inner membrane serves to illustrate how this process can activate the protein, by bringing it into close proximity with its substrate. Proteins can also be induced to translocate to a multiprotein complex known as the proteasome, which degrades cellular proteins (providing, as well as a mechanism of protein removal, a source of peptides for presentation by MHC molecules). An interesting example of this process involves the human papilloma viral (HPV) protein E6 [17]. The function of E6 appears to be to direct the tumor suppressor protein p53, produced by the host cell, to the proteasome, where it is proteolytically degraded.

This seems to be an important factor in the ability of high-risk strains of HPV to cause cervical cancer. E6 binds to a cellular protein, the E6-associated protein (E6AP); the complex then binds to p53 and directs it to a set of enzymes involved in the ubiquitination of lysine sidechains (Fig. 11). Ubiquitinated proteins are 'tagged' for degradation by the proteasome. E6 is thus acting as a heterodimerizing agent, bringing p53 close to the machinery that identifies proteins for removal. The structural aspects of this overall process have not yet been studied in detail, but the overall process is another clear illustration of a proximity effect. Proteins do not wander stochastically on the way to their fateful encounter with the proteasome, but are induced to become targets for it by the action of molecular dimerizers.

Synthetic dimerizers: new opportunities for intervention

Our familarity with the concepts of conformational change induced by allosteric agents has facilitated the discovery of

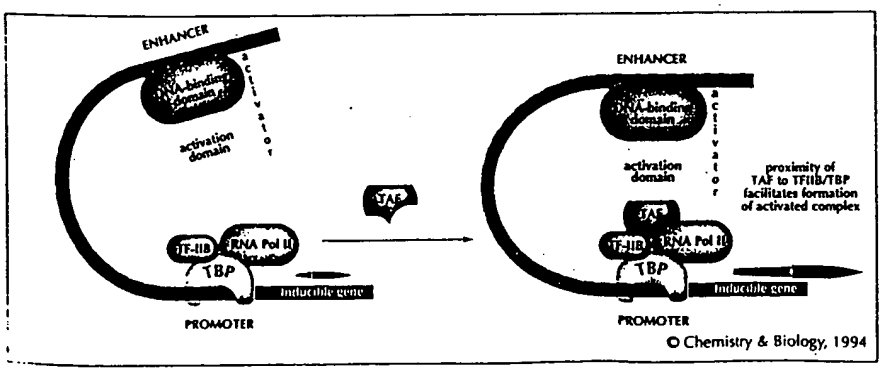


Fig. 10. Mechanism of transcriptional activation by transcriptional activators involving the recruitment of general transcription factors to a promoter sequence.

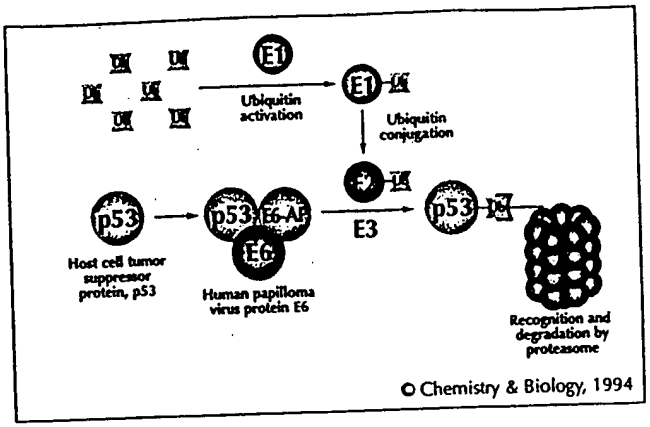


Fig. 11. Human papilloma virus-mediated ubiquitination of the tumor suppressor p53. Ubiquitination of the target protein occurs through initial conjugation of the ubiquitin-activating enzyme (E1) with the carboxy-terminal glycine of ubiquitin (Ub) via a cysteine residue, followed by a cysteine residue transfer from E1 to ubiquitin-conjugating enzyme (E2/E3). E6 and the E6-associated protein (E6-AP) bind specifically to the tumor suppressor p53, and, acting as an adaptor, work with E2 to facilitate the degradation of p53 via the ubiquitination pathway.

synthetic versions of these molecules. Many examples of synthetic allosteric agents can be found in medicine, and some of these have facilitated the study of basic cellular mechanisms. By recognizing the prominent role of natural dimerizers in cell biology, the properties of organic equivalents can now also be readily and reliably predicted. Formulating the structures of CIDs and synthesizing them will be a new, and worthwhile, challenge. Nature has provided clear illustrations of the feasibility of this approach in the form of the natural products CsA, FK506, and rapamycin. These low-molecular-weight, organic heterodimerizers are monomeric molecules equipped with two distinct protein-binding surfaces. One approach to constructing a designed CID involved converting the monomeric heterodimerizer FK506 into the dimeric homodimerizer FK1012 by eliminating one of FK506's protein-binding surfaces and replacing it with an element that crosslinked the modified natural product [18,19].

It is easy to imagine the mixing and matching of different protein-binding surfaces using synthetic organic chemistry, to create new dimerizers with tailor-made properties. Since protein dimerizers simply create a high local concentration of a particular protein at a particular cellular location, their actions will not require the geometric precision associated with allosteric agents. This fact, combined with our increasing ability to generate specific protein-binding surfaces either by screening or by design, suggests that synthetic dimerizers might be even more readily accessible than the classical allosteric agents.

References

1. Linder, M.E. & Gilman, A.G. (1992). G proteins. *Sci. Am.* **267**, 36–43.
2. Cunningham, B.C., Ultsch, M., de Vos, A.M., Mulkerrin, M.G., Clauser, K.R. & Wells, J.A. (1991). Dimerization of the extracellular domain of the human growth hormone receptor by a single hormone molecule. *Science* **254**, 821–825.
3. de Vos, A., Ultsch, M. & Kossiakoff, A.A. (1992). Human growth hormone and extracellular domain of its receptor: crystal structure of the complex. *Science* **255**, 306–312.
4. Chou, K.-C., Carlacci, L., Maggiora, G.M., Parodi, L.A. & Schulz, M. W. (1992). An energy-based approach to packing the 7-helix bundle of bacteriorhodopsin. *Prot. Sci.* **1**, 810–827.
5. Zhang, R.-G., Joachimiak, A., Lawson, C.L., Schevitz, R.W., Otwinowski, Z. & Sigler, P.B. (1987). The crystal structure of *trp* aporepressor at 1.8 Å shows how binding tryptophan enhances DNA affinity. *Nature* **327**, 591–597.
6. Otwinowski, Z., *et al.*, & Sigler, P.B. (1988). Crystal structure of the *TRP* repressor/operator complex at atomic resolution. *Nature* **335**, 321–329.
7. Merluzzi, VJ., *et al.*, & Sullivan, J.L. (1990). Inhibition of HIV-1 replication by a nonnucleoside reverse transcriptase inhibitor. *Science* **250**, 1411–1413.
8. Kohlstaedt, L.A., Wang, J., Friedman, J.M., Rice, P.A. & Steitz, T.A. (1992). Crystal structure at 3.5 Å resolution of HIV-1 reverse transcriptase complexed with an inhibitor. *Science* **256**, 1783–1790.
9. Pawson, T. & Schlessinger, J. (1993). SH2 and SH3 domains. *Curr. Biol.* **3**, 434–442.
10. Waksman, G., Shoelson, S.E., Pant, N., Cowburn, D. & Kuriyan, J. (1993). Binding of a high affinity phosphotyrosyl peptide to the Src SH2 domain: crystal structures of the complexed and peptide-free forms. *Cell* **72**, 779–790.
11. Aronheim, A., Engleberg, D., Li, N., Al-Alawi, N., Schlessinger, J. & Karin, M. (1994). Membrane targeting of the nucleotide exchange factor Sos is sufficient for activating the Ras signaling pathway. *Cell* **78**, 949–961.
12. Liu, J., Farmer, J.D., Lane, W.S., Friedman, J., Weissman, I. & Schreiber, S.L. (1991). Calcineurin is a common target of cyclophilin–cyclosporin A and FKBP–FK506 Complexes. *Cell* **66**, 807–815.
13. Brown, E. J., *et al.*, & Schreiber, S. L. (1994). A mammalian protein targeted by G1-arresting rapamycin–receptor complex. *Nature* **369**, 756–758.
14. Sabatini, D.M., Erdjument-Bromage, H., Lui, M., Tempst, P. & Snyder, S.H. (1994) RAFT1: a mammalian protein that binds to FKBP12 in a rapamycin-dependent fashion and is homologous to yeast TORs. *Cell* **78**, 35–43.
15. Stern, L.J. & Wiley, D.C. (1994). Antigenic peptide binding by Class I and Class II histocompatibility proteins. *Structure* **2**, 245–251.
16. Tijan, R. & Maniatis, T. (1994). Transcriptional activation: a complex puzzle with few easy pieces. *Cell* **77**, 5–8.
17. Scheffner, M., Huibregtse, J.M., Vierstra, R.D. & Howley, P.M. (1993). The HPV-16 E6 and E6-AP complex functions as a ubiquitin-protein ligase in the ubiquitination of p53. *Cell* **75**, 495–505.
18. Spencer, D.M., Wandless, T.J., Schreiber, S.L. & Crabtree, G.R. (1993). Controlling signal transduction with synthetic ligands. *Science* **262**, 1019–1024.
19. Pruschy, M., Spencer, D.M., Kapoor, T.M., Miyake, H., Crabtree, G. R. & Schreiber, S.L. (1994). Mechanistic studies of a signaling pathway activated by the organic dimerizer FK1012. *Chemistry & Biology* **1**, 163-172.
20. Nicholls, A., Sharp, K.A. & Honig, B. (1991). *Prot. Struct. Funct. Genet.* **11**, 282–293.

David J Austin and Stuart L Schreiber, Howard Hughes Medical Institute, Department of Chemistry, Harvard University, 12 Oxford Street, Cambridge, MA 02138, USA and Gerald R Crabtree, Howard Hughes Medical Institute, Departments of Developmental Biology & Pathology, Stanford University School of Medicine, Stanford, CA 94305, USA.

EXHIBIT E

JOURNAL OF
MEDICINAL
CHEMISTRY

Volume 37, Number 10 May 13, 1994

Perspective

Applications of Combinatorial Technologies to Drug Discovery. 2. Combinatorial Organic Synthesis, Library Screening Strategies, and Future Directions[1]

Eric M. Gordon,* Ronald W. Barrett, William J. Dower, Stephen P. A. Fodor, and Mark A. Gallop

Affymax Research Institute, 4001 Miranda Avenue, Palo Alto, California 94304

Received March 1, 1994

Biographies

Ronald W. Barrett received his Ph.D. in Pharmacology from Rutgers University in 1983 and did postdoctoral work at the Addiction Research Foundation in Palo Alto, CA, in the field of opiate receptors. From 1986 to 1989, he worked in the Neuroscience Research Area at Abbott Laboratories in Abbott Park, IL. He joined Affymax in 1989 where he established the Receptor Pharmacology Group. Dr. Barrett is presently Vice President and Director of Receptor Pharmacology.

William J. Dower received his Ph.D. in Biology from the University of California, San Diego, and did postdoctoral work at Stanford University studying steroid control of specific mRNA metabolism and the amplification of genes mediated by chemotherapeutic agents. He joined Bio-Rad Laboratories in 1984 where as a member of the newly formed molecular biology unit, he introduced the electroporation technique for the high-efficiency transformation of bacteria. He joined Affymax in 1989 to establish a molecular biology research group. Dr. Dower is Director of Molecular Biology at Affymax.

Stephen P. A. Fodor received his Ph.D. in Chemistry at Princeton University and was an NIH postdoctoral fellow in Chemistry at the University of California, Berkeley. He joined Affymax in 1989 as a Staff Scientist in Optical Technologies, becoming Director of Physical Sciences. Dr. Fodor's group led the development of new technologies, merging photolithography with combinatorial solid-phase chemistry. In 1993 he joined Affymetrix as Scientific Director, where he is currently using oligonucleotide arrays to study a variety of DNA molecular recognition processes.

Mark A. Gallop received his Ph.D. from the University of Cambridge, England, for research in the area of organo transition metal chemistry and did postdoctoral work in the laboratories of Peter G. Schultz at the University of California, Berkeley. He joined Affymax in 1990 and is currently a Senior Scientist and leader of the Combinatorial Chemistry Group.

Eric M. Gordon received his Ph.D. in 1973 from the University of Wisconsin—Madison and did postdoctoral work at Yale University. His research interests include the rational design of enzyme inhibitors, the chemistry of amino acids, peptides, and natural products, and combinatorial chemistry. Dr. Gordon has authored 150 papers and U.S. patents in these areas. From 1974 to 1992 he worked at the Squibb Institute for Medical Research (subsequently the Bristol-Myers Squibb Pharmaceutical Research Institute) in Princeton, NJ, most recently as a Director of Medicinal Chemistry. In 1992, he joined the Affymax Research Institute where he is currently Vice President of Research and Director of Chemistry.

A. Combinatorial Organic Synthesis

The notion of creating huge, searchable libraries of small organic molecules is unprecedented in medicinal chemistry, and the possibility of doing so has recently captured the imagination of the drug-discovery community. The conventional paradigm of small molecule lead development, in which a compound undergoes many rounds of individualized, hand-crafted modification and biological testing en route to drug candidacy, will likely be dramatically accelerated by the application of combinatorial chemistry technologies to mass-produce and evaluate lead analogs. The ability to harness molecular diversity techniques as tools for lead discovery offers an unparalleled opportunity for medicinal chemistry to expand the breadth and scope of molecular structures that may be screened for biological activity. Widespread availability of collections of highly diverse small-molecule libraries should provide an opportunity to assess the impact of combinatorial organic synthesis on new-lead discovery. In this section of part 2, some of the issues which confront the practitioner of combinatorial organic synthesis, as they relate to the problems of molecular recognition in general and medicinal chemistry in particular, will be analyzed.

Issues in Practicing Combinatorial Organic Synthesis

Combinatorial organic synthesis (COS) presents somewhat of an intellectual inversion of the past 50 years of synthetic organic chemistry. The chemist of the Woodwardian era was interested in a masterful and carefully

0022-2623/94/1837-1385$04.50/0

plotted natural product synthesis of a complex entity of known structure. Reactions were more often employed or developed to solve specific challenges rather than to provide generic methodologies. Rigorous control of reaction pathways, stereochemistry, and regiochemistry, and the exclusion of all but the desired diastereomer were obligatory in a faithful rendering of the technique. In contradistinction to natural product total synthesis, rather than generating a single, specific entity, the goals of COS are to create *populations* of molecular structures. Rather than exercising complete control, the combinatorial chemist, while maintaining high reaction efficiency and relative reactive compatibility, may actually seek to create situations and apply strategies in which stereochemical/regiochemical control is relaxed. This must be achieved while remaining cognizant of the impact these factors may have on the stoichiometry of the resulting libary and its design and ultimate use. Hence, the combinatorial chemist seeks to apply a series of Woodwardian reactions (reliable, high yielding) that operate generically on a diverse set of building blocks to provide a multitude of related products.

Criteria for Library Design. The primary objectives of producing small-molecule libraries by COS are to provide collections of compounds suitable for both drug-discovery screening and drug-development optimization. When complete, the combinatorial drug-discovery exercise should have created a stable population of low molecular weight entities, free of reactive and toxicity-causing functionality. While a paramount medicinal chemistry design criteria for small-molecule-libary construction might be that the *products* of diversity generation (individual library members) should "look" like drug leads, of still greater importance is that the library actually contains compounds capable of interacting at some detectable level with the biological target of interest. When small-molecule leads for a target have been previously defined (e.g., benzodiazepine ligands for a peptide or other G-protein-coupled receptor, transition-state inhibitors for a protease), the notion of searching for more potent derivatives among libraries combinatorially enriched in specific pharmacophore analogs is an obvious tactic to pursue. However, as the universe of well-defined macromolecular drug targets continues to expand through the impact of molecular cloning, the problem of identifying new pharmacophores capable of modulating the various interactions of peptides, proteins, carbohydrates, oligonucleotides, or lipids at these sites will also be intensified.

Will "rules" about the types of libraries that may prove generally useful in ligand discovery be discovered? Although the field of molecular diversity has not yet matured to the point where substantial insight into this question is forthcoming, it is intuitively obvious that small-molecule libraries, whose members structurally resemble historical leads, should provide a fertile reservoir of potential molecular diversity. Tangential to this, natural products aside, numerous historical drug leads were derived simply because synthetic routes to these molecules were readily available. It is likely that early-stage COS will be limited by applicable chemistry and that this will necessarily focus work toward traditional leads, whose syntheses are known and well-documented.

The successful identification from recombinant libraries of L-amino acid-based peptide ligands that inhibit protein–peptide, protein–protein, and protein–carbohydrate interactions suggests there is broad utility in screening large libraries of peptidic compounds. It remains to be seen whether collections of other random molecular structures that are quantitatively as diverse as existing peptide libraries prove in *de novo* ligand discovery to include the "pharmacophores of the future".

Ligand rigidity may be another important parameter to consider in the course of library design. The incorporation of conformational constraints into flexible lead molecules has emerged as a powerful strategy to enhance ligand potency and/or selectivity, particularly in the field of peptidomimetic medicinal chemistry.[2-10] Nevertheless, with regard to library design, conformational restriction may act as a two-edged sword: an inappropriate constraint is likely to abrogate the modest but perhaps detectable activity of a more flexible analog, which could, in a secondary library, be systematically constrained. From the point of view of random screening, it remains to be determined whether useful leads will arise more frequently from libraries of rigidified or flexible structures. Data from the evaluation of cyclic peptide libraries in both synthetic and recombinant systems may provide some important insights into this issue. A number of methods have been recently described for on-resin cyclization of peptides through both main-chain and side-chain functional groups.[6-10,68] At present, a portfolio of libraries containing both conformationally rigid and relaxed molecular diversity seems most appropriate. A longer range solution might be to moderate the high risk of conformational restriction by creating very large populations of semirigid molecular arrays, comprising structural families that collectively sample as completely as possible all regions of conformational space.

Characterization. The usual measures of evaluating success in organic synthesis may lose meaning in COS. The classical notions of such fundamental concepts as purity/homogeneity, yield, exact product structure, relative and absolute stereochemical control, specific physical properties are less relevant when applied to a broad population of molecules (of course they may become quite relevant as individuals emerge from a selection process). Additionally, the analytical mainstays of the synthetic organic chemist, such as NMR and IR, may become obviated. The NMR spectrum of a 10 000-component library mixture is not diagnostic. The loss of these powerful tools requires that compensating technologies be developed. A major dilemma of COS is the difficulty of confirming the degree to which the expected chemistry has proceeded on the entire population of substrate molecules. Several groups have recently reported on the use of electrospray mass spectrometry as a technique for evaluating the bulk composition of diverse peptide libaries.[11,12] Gross synthetic discrepancies, such as incomplete protecting group removal, may be detectable by mass analysis, providing an opportunity to optimize the library synthesis protocols. In the characterization of combinatorial products, the presence of "byproducts" (in COS, unexpected products), combined with the difficulty of detecting these compounds, will cause problems if one mistakenly concludes that a screening hit is the expected product. This section will conclude by offering a potential solution to this problem.

Efficiency/Automation. Among the chemical criteria relevant for small-molecule-library design is the efficiency of diversity creation. The assembly of most small molecules reduces to the intercombination of only three to

five building blocks of molecular weight ~150 each. Synthetic reactions capable of combining numerous building blocks simultaneously constitute a highly efficient form of diversity generation. Thus the Ugi four-component reaction has a high combinatorial efficiency since building blocks of four families (amines, carbonyl compounds, isocyanides, and suitable acid components) are linked simultaneously to afford α-amino acid derivatives. In contrast, peptide chemistry traditionally links two building blocks at a time. In both the broad screening and the lead analoging modes, a longer range question pertains to the ability of the chemistry to eventually be automated. Once the key decisions and overall strategy have been determined, much of the actual chemistry is repetitive in nature. Machines will continue to be constructed to capitalize on this and libraries will be assembled under computer control.[13-15]

Quantity and Quality of Diversity. While the "quantity of diversity" that is experimentally accessible can be dictated by the number of building blocks in the basis set and by the number of synthetic operations applied, or able to be applied (see part 1[1]), the practical limitations on library size are most generally imposed by the format within which the diversity is created and evaluated. A small number of building blocks subjected to many synthetic steps will yield high (numerical) diversity; however the products of these reactions may be relatively large molecules, not well-suited for lead development as traditionally administered therapeutics. Thus, as the combinatorial process proceeds, an opportunity window may exist in which the bulk of the library possesses properties which standard medicinal chemistry usually seeks in small-molecule drug discovery (MW < 700, solubility, etc.). Continued application of the combinatorial process will lead to product libraries containing larger molecules (composed of more building blocks) wherein the individual library members have "outgrown" the classical criteria of a lead-drug molecule.

In surveying the historical landscape of drug discovery, there are particular pharmacophores or structural arrays which periodically surface far in excess of random chance (benzodiazepines, β-lactams, imidazoles, phenethylamines, etc.).† A review of recent successes in the era of "rational drug design" suggests that certain molecular concatenations—protein turn mimetics, conformationally restricted amino acids, transition-state analogs, dipeptide isosteres, molecular scaffolds, designed elements for enzyme inhibition—are often found in the medicinal chemistry of lead compound development. In consideration of the molecular structures which have left their mark on modern medicinal chemistry, one might conclude that the drug-discovery process is impacted not only by the sheer quantity of diversity surveyed, but additionally by the more subjective "quality" of diversity that is evaluated. Different organizations and individuals will certainly bring a wide variety of criteria to the subjective appraisal process, depending on style, experience, and bias. It may be speculated that the quality of diversity will be influenced by the sophistication of the building blocks originally introduced into the combining system (library bias on the part of the medicinal chemist) and the extent to which molecular substructures of the building blocks can be assembled in diverse, spatial (3-D) relationships. Thus the collected expertise of medicinal chemical knowledge may be used to "hyperevolve" or "bias" the libary by the planned introduction of commonly evolved elements; these elements are "retrocombinatorial synthons" of many known bioactive classes. Thus the building block basis set must be judiciously chosen and carefully attuned to the collected knowledge historically amassed in drug discovery.

† An interesting aside regarding these important substructures is that development of "generic" syntheses of key pharmacophores ultimately enabled facile generation of many analogs. Concurrently or subsequently, diverse biological activities were found among these compound classes. In a sense, this is suggestive of combinatorial chemistry, except the crucial molecules were made serially rather than in a parallel/combinatorial high throughput fashion.

Issues in the Selection of Building-Block Sets

The acquisition of a building-block library can be a major time and resource investment, and the eventual decision of which type of chemical building blocks to utilize places limits on the universe of structural diversity which ultimately can be explored. Depending on the specific objective, important building-block criteria include the availability of a large number of diverse, fairly complex, easily accessible starting materials. These may be either commercially available or prepared in a few steps from commercial materials. Members of a building block set should reflect a broad array of physicochemical properties, functionality, charge, conformation, etc. Building blocks may be chiral, achiral, or racemic. Certain building-block families have what may be termed a high "combinatorial potential". This relates to the high density per carbon atom of reactive functionality which can participate in new covalent combinations. For example, monosaccharides have high combinatorial potential since the high density of available hydroxyl groups leads to many potential connecting permutations. In addition to polymer formation, the high combinatorial potential of such types of building blocks may also be exploited as scaffolds for the generation of diversity (*vide infra*).

Synthetic Strategy

An important strategic element in combinatorial library synthesis is the degree of reliability of the ligand synthesis chemistry. What is the likelihood of general synthetic success with a particular reaction? The nature of combinatorial reactions, which must proceed in the face of a broad range of functionality on a multitude of substrates and where the products are difficult to analyze individually, demands that, in selection of synthetic methodologies, greater weight must be given to reaction sequences with reliable, predictable outcomes. A more subtle question revolves around the number of synthetic options available in the course of diversity generation. For example, a synthetic strategy structured in such a way that, as the process proceeded, new combinatorial possibilities opened up, would be preferable to having options narrow, especially if the goal was generating a maximum of structural diversity.

As previously noted, there are two distinct themes that must be considered for the successful application of combinatorial technologies to ligand discovery and optimization, *viz.* broad-based screening and directed chemical analoging. The issues underlying conceptual design, as well as the synthetic strategies utilized in construction of these different classes of libraries, are noteworthy and are summarized in Figure 1. Building block requirements for undertaking broad and narrow diversity searches differ markedly. The search for an initial lead molecule may be essentially a random screening exercise, where the em-

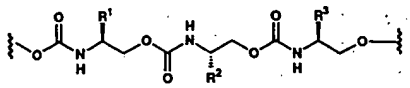
Broad Screening	Chemical Analoging/Optimization
huge size library	modest size library
broadest structural diversity	relatively narrow structural diversity
no special initial structure goal	specific structural goal
any building blocks	specific retrocombinatorial building blocks
undefined order of reaction	specific order of combination
flexible synthetic strategy	well defined synthetic strategy
site of tether not crucial	tether crucial-build in redundancy
ligand possibly uncouplable	ligand should be releasable
single selection evolution	cumulative selection evolution

Figure 1. Combinatorial chemistry: comparison of two major themes.

phasis is on exposing a macromolecular drug target to the maximum possible structural diversity. The objective is to identify a ligand of significant affinity for the target, the exact ligand structure and its detailed characteristics at this point are not relevant: in fact any molecule will do. An approach to generating highly diverse libraries for use in medicinal chemistry might favor using building blocks which have distinguished themselves by appearing frequently in previous active leads (e.g., statine, hydroxyethylamines, Freidinger lactams[16]). On the other hand, once a lead is available, most drug discovery proceeds through a series of evolutions (optimizations) in order to meet a set of predetermined criteria. Since specific structural types are sought, searching in a very broad pool of diversity (as above) is unlikely to be successful (actually it could uncover a new lead but is less likely to optimize an existing one). Ideally, what is required in this type of diversity-generating strategy is to "explode" around the known lead, i.e., to create as highly diverse population as possible that bears close structural resemblance to the original hit, followed by a selection for desired criteria.

Clearly the subunits which lead to predetermined structures must be quite specific: from where should building blocks for known structural classes of pharmacophores arise? The answer, as in organic synthesis, lies in a retrosynthetic analysis or what we may term a *retrocombinatorial* approach to building-block selection. Lead structures should be retrosynthetically dissected in the maximum number of ways and upon these various possibilities imposed the needs of performing combinatorial chemistry. Inspection of the retrosynthetic tree invites the following key questions: By which modes of forward synthesis are the most building blocks available or obtainable? If the synthesis is allowed to proceed by that course, what is the scope and degree of reliability of the necessary reactions? Extending this line of reasoning should permit the maximum leverage to be applied combinatorially.

A common feature of both paradigms is likely to be a reliance on solid-phase-synthesis methods to facilitate the assembly of combinatorial libraries. Synthesis on a polymeric support greatly simplifies the problem of product isolation from reaction mixtures and also facilitates the partitioning of products into multiple aliquots for subsequent chemical elaboration. Moreover, the opportunity exists to take advantage of the support-tethered diversity in the design of convenient receptor binding assays for library evaluation. While there has been a long tradition of polymer-supported organic chemistry,[17-20] it is only in the areas of peptide and



Polycarbamates

Figure 2. Structure of a synthetic oligocarbamate prototype.

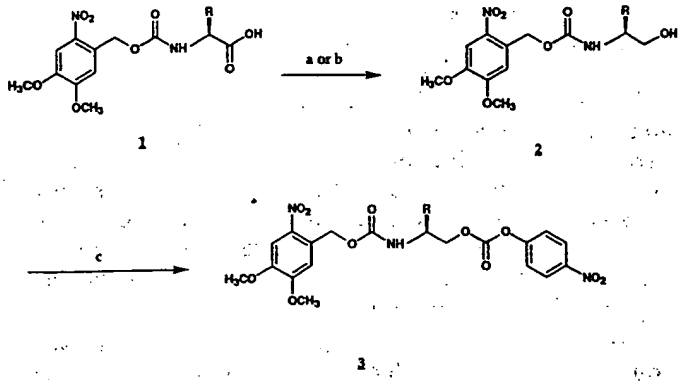


Figure 3. Synthesis of N-protected nitrophenyl carbonate monomers. Key: (a) BH_3, THF; (b) DCC, methylene chloride, *N*-hydroxysuccinimide, HOBt; then sodium borohydride, ethanol; (c) *p*-nitrophenyl chloroformate, methylene chloride, pyridine.

oligonucleotide synthesis that solid-supported chemistry has truly been optimized and become common-place. The advent of combinatorial organic synthesis will undoubtedly signal a renaissance in solid-phase organic chemistry, as workers attempt to adapt well-characterized homogeneous reactions to reliable solid-supported protocols.

Progress to Date: Synthetic Polymeric Diversity

The design and synthesis of novel synthetic monomers which, when assembled in a combinatorial fashion, could yield relatively low molecular weight polymeric materials is an approach that is well-suited to diversity generation and evaluation. Combinations of such monomers could lead to substances with novel backbones, possibly possessing desirable properties, such as metabolic stability, enhanced pharmacokinetic profiles, and cell and membrane permeability. Identification of these and other potentially modifiable parameters in such systems could facilitate drug discovery.

Schultz and co-workers have reported the synthesis of a library of oligocarbamates starting from a basis set of chiral aminocarbonates[21] (Figure 2). The monomeric units were readily obtained by the modification of amino acids via the intermediacy of the corresponding chiral amino alcohols (see Figure 3). The resulting nitrophenyl carbonate building blocks (*3*) were stable for several months at room temperature.

Oligocarbamates were synthesized on a solid support by deprotection of a resin-bound amine, protected with either the base-labile Fmoc or photolabile nitroveratryloxycarbonyl (Nvoc) group, followed by treatment with a nitrophenyl carbonate of type 3. The deprotection/coupling cycle was repeated until an oligocarbamate of the desired length was attained (seven or eight cycles). Overall coupling yields were greater than 99% per step. Side-chain deprotection followed by resin cleavage afforded the desired oligocarbamates (Figure 4).

The VLSIPS photolithographic chip format, previously employed for oligopeptide synthesis, was used in the construction a spatially-addressable oligocarbamate library of 256 members. An anti-carbamate monoclonal antibody served as a model receptor for screening against this array. Antibody:oligocarbamate complexes were

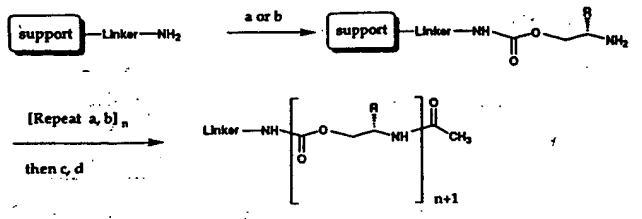


Figure 4. Solid-phase synthesis of oligocarbamates. Key: (a) nitrophenyl carbonate monomer, HOBt, diisopropylethylamine, NMP; (b) piperidine, NMP or *hν*; (c) acetic anhydride, NMP; (d) TFA, triethylsilane.

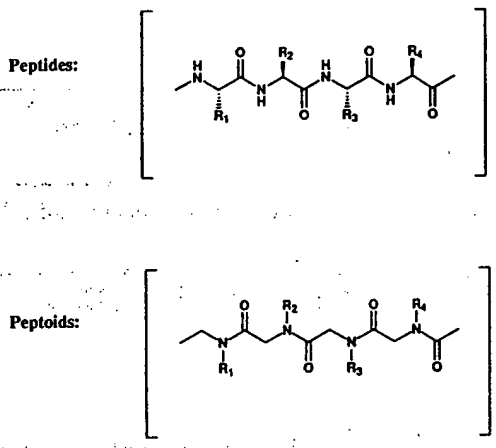


Figure 5. Comparison of peptide and peptoid backbones.

detected by treatment of the chip with a fluorescein-conjugated secondary antibody, followed by analysis using scanning epifluorescence microscopy. Because the location and structure of each different library member is defined by the synthetic strategy (binary masking) used in this technique, the necessity of sequencing the products is obviated. The binding activities of putative hits were confirmed by conventional assays using authentic material prepared by independent synthesis. A preliminary evaluation of the physicochemical properties of oligocarbamate molecules indicate that they are more hydrophobic than the corresponding peptide homologs, and their expected resistance to several proteases was confirmed.

Another type of synthetic polymeric diversity has been developed by Simon *et al.*[22,23] Through a variety of preparative routes, this group created a basis set of monomeric N-substituted glycine units, each bearing a nitrogen substituent similar to those of the natural α-amino acid side chains. The formal polymerization of these monomers results in a class of polymeric diversity which these workers have termed "peptoids" (Figure 5). Peptoids may be synthesized either manually or robotically following either a "full monomer" oligomer synthesis or via a "submonomer" synthesis, as reported by Zuckermann *et al.*[24] and illustrated in Figure 6. Various biological activities have been established for specific peptoid sequences, including inhibition of α-amylase and the hepatitis A virus 3C protease, binding to the tat RNA of HIV[22], and antagonism at the α_1-adrenergic receptor.[25] The peptoid approach to diversity generation has been extended to the preparation of encoded combinatorial libraries, in which natural amino acids code for the structure of the peptoid chain[26] (see part 1[1] and Figure 7).

An important variant of the synthetic polymeric diversity approach is directed toward construction of a chemical library in which the peptidyl backbone is conserved but a dipeptide unit is replaced at specific

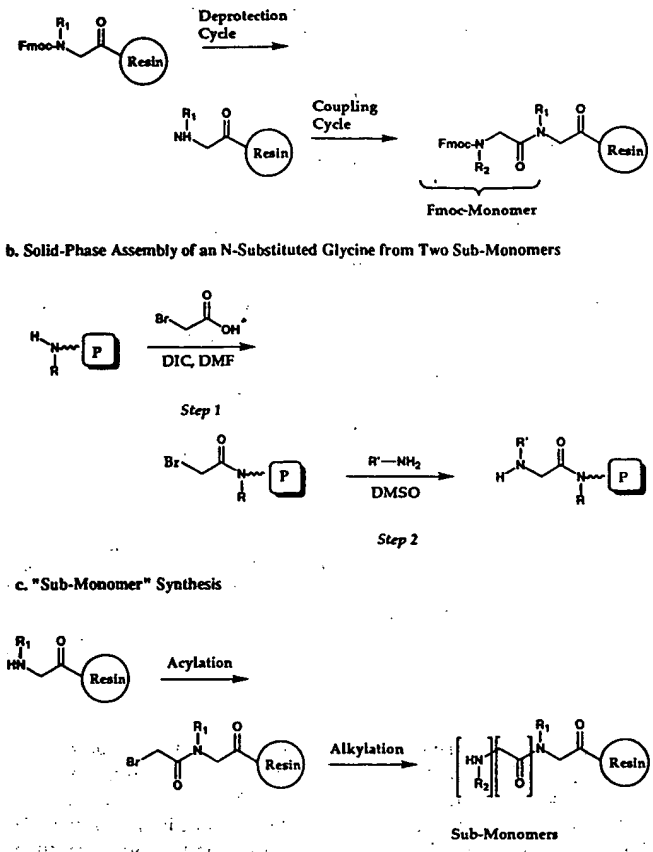


Figure 6. Synthesis of peptoids.

positions by a phosphonate dipeptide surrogate (see Figure 8). Such phosphonate pharmacophores are well-known as transition-state analogs for amide bond cleavage and have found wide usage in the inhibition of metalloproteases.[27-29] Campbell has described methods for the solid-phase synthesis of peptidylphosphonates that are compatible with the Fmoc/^{t}Bu protecting group strategy of standard peptide synthesis.[30] The key reaction step is formation of the phosphonate ester bond, which is achieved via a modified Mitsunobu condensation (Figure 9). Precursor lactic acid and protected amino phosphonate building blocks are prepared as shown in Figure 10.

When this process is applied to the combinatorial synthesis of peptidylphosphonates, the diversity product will be a metalloprotease enzyme inhibitor library. Enzyme-inhibitor libraries of this type and those focusing on other known inhibitory pharmacophores (thiols, hydroxamates, carboxyalkyldipeptides, etc.) may prove to be important tools in rapidly profiling novel proteases and for determining which pharmacophores are most effective at their inhibition. Using this knowledge, secondary inhibitory libraries can be constructed to optimize original leads. Through such a process it may be possible to dramatically accelerate the process of finding high-affinity enzyme-inhibitor ligands.

Another interesting type of polymeric diversity based upon a vinylogous polypeptide backbone has recently been reported by Hagihara *et al.*,[31] in which introduction of a trans olefinic linkage between the α-carbon and the carbonyl group of various amino acids is generalized. Additionally, Smith and colleagues have synthesized a non-amide polymer of (3,5)-linked pyrrolin-4-one oligomers which mimic the β-strand conformation of a normal peptide chain[32] (see Figure 11).

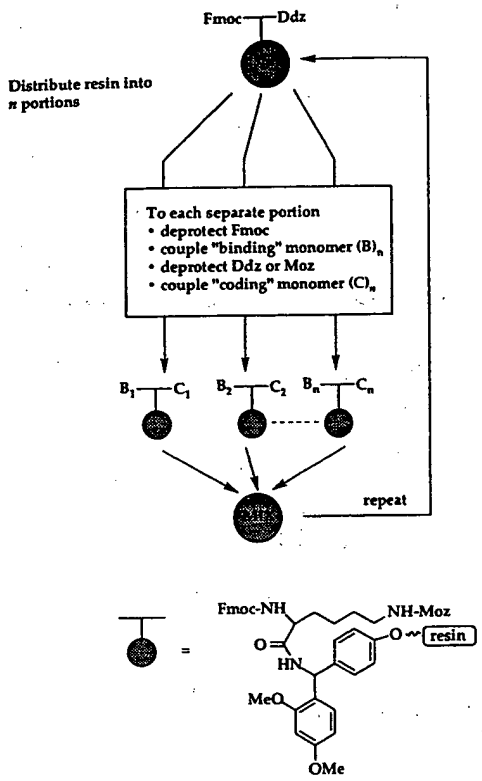


Figure 7. The synthesis of an encoded library consists of the following steps: (1) A bifunctional linker containing two orthogonally protected points of extension (Nα-Fmoc-Nϵ-Moz-Lys-OH) is attached to polystyrene resin via an acid-labile linker. (2) The solid support is divided into n equal portions at a mixture position. (3) A unique Nα-Fmoc-protected non-natural monomer (B) is coupled to the "binding" strand. (4) A series of Nα-Ddz-protected L-amino acids (C) are then coupled to the "coding" strand. (5) The solid supports are recombined.

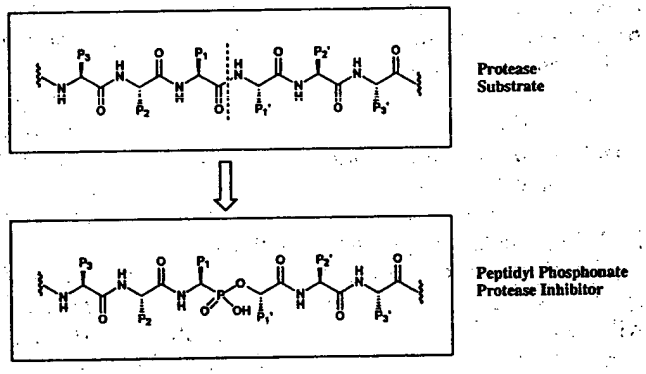


Figure 8. Peptidyl phosphonates.

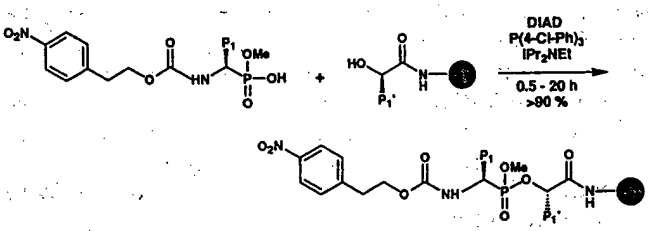


Figure 9. Solid-phase peptidyl phosphonate synthesis (SPPPS).

Nonpolymeric, Small-Molecule Diversity. The majority of chemical diversity generation discussed above concerns the preparation of linear molecules, in which the target structures are unambiguously specified by the order of building-block addition. In contrast, the great preponderance of organic synthesis proceeds rather differently, wherein building blocks interlock to give rise to

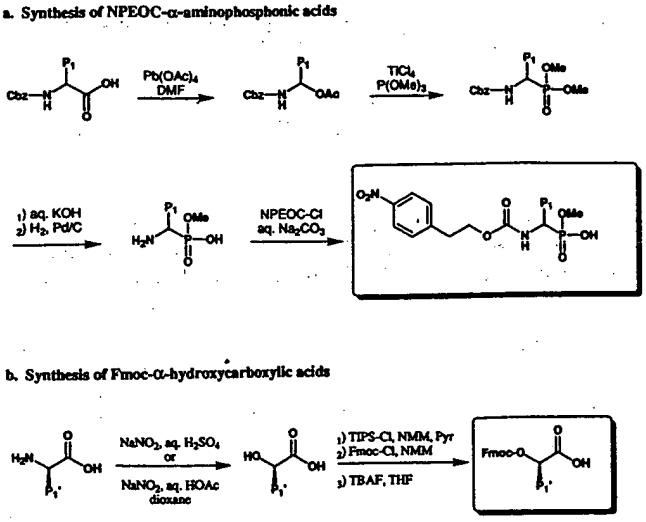


Figure 10. Synthesis of peptidyl phosphonate building blocks.

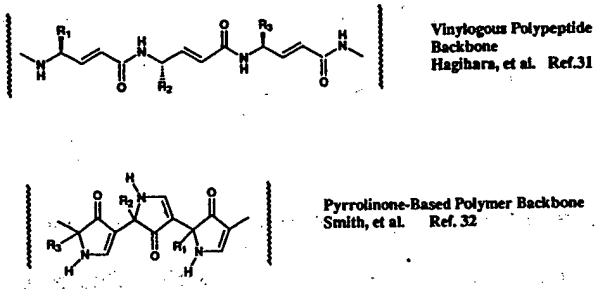


Figure 11. Novel polymeric backbones.

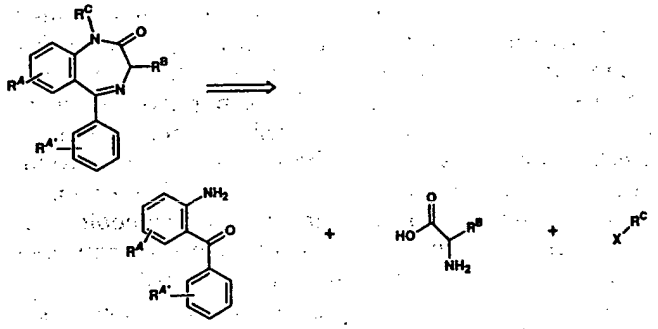


Figure 12. Components of a benzodiazepine library.

nonpolymeric, three-dimensional arrays. The recent seminal work of Ellman on the solid-phase synthesis of 1,4-benzodiazepines lays the groundwork for creation of a small-molecule library of one of medicinal chemistry's most notable pharmacophores and represents one of the first examples of the application of combinatorial organic synthesis to nonpolymeric organic compounds.[33]

The benzodiazepines are synthesized on a solid support by the connection of three building blocks, each of different chemical families (Figure 12). Following the attachment of 2-aminobenzophenone hydroxy or carboxy derivatives to the support using an acid-cleavable linker, [(4-hydroxymethyl)phenoxyacetic acid], the N-protecting group is deblocked (piperidine/DMF), and the weakly nucleophilic amine is acylated with an α-Fmoc-protected amino acid fluoride, using 4-methyl-2,6-di-*tert*-butylpyridine as an acid scavenger (Figure 13). Fmoc deprotection, followed by treatment with 5% acetic acid in DMF, causes the general cyclization to the intermediate lactam. Capitalizing on the ability of lithiated 5-(phenylmethyl)-2-oxazolidinone to selectively deprotonate the anilide NH,

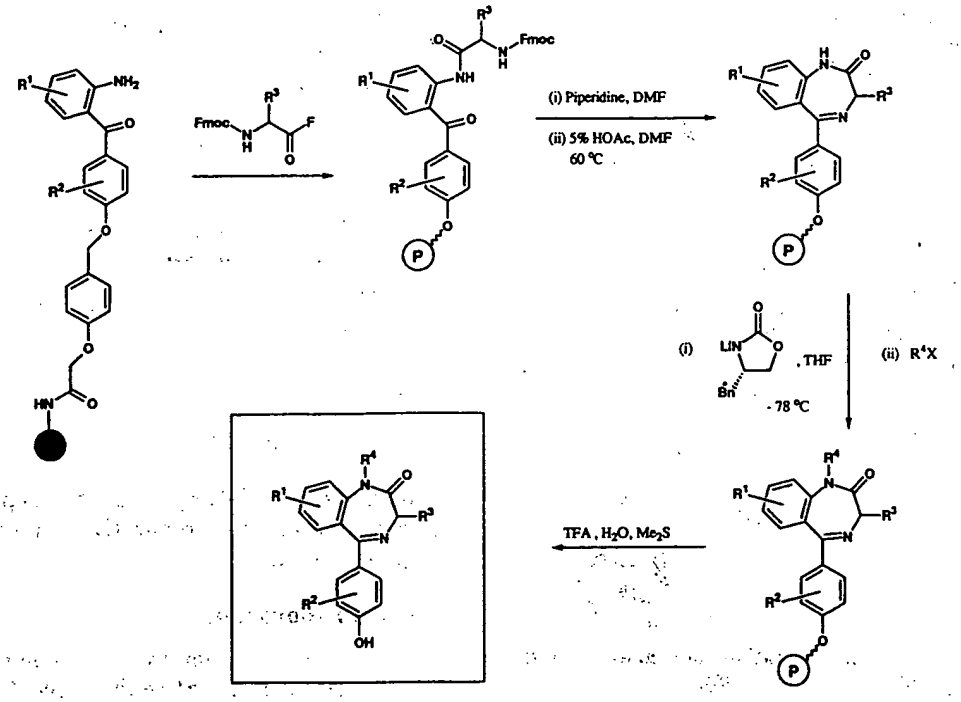


Figure 13. Combinatorial synthesis of the benzodiazepine pharmacophore.

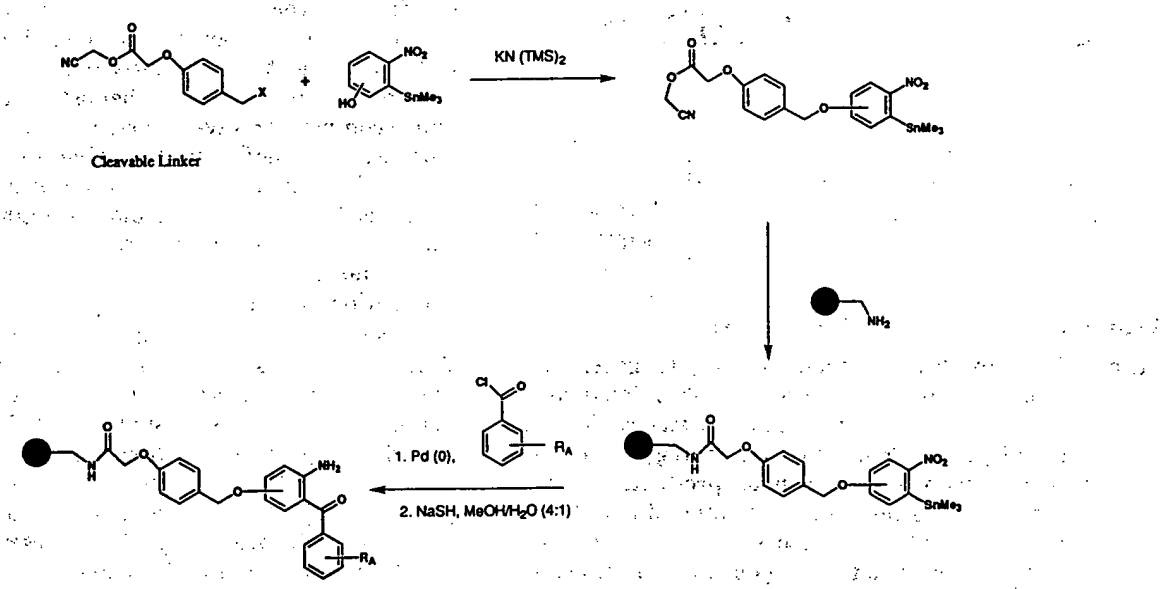


Figure 14. Synthesis of 2-aminobenzophenone derivatives on a solid support.

alkylation was achieved with a variety of alkylating agents. Aqueous acid cleaves the new benzodiazepine from the support in very high overall yields. The integrity of the chiral center was confirmed by a racemization test using chiral HPLC.

One of the limiting features of applying the above scheme to combinatorial library construction is that, though many alkylating agents and amino acid building blocks are commercially available, there is not a ready supply of appropriately functionalized 2-aminobenzophenones. Ellman addressed this problem directly by creating a general method for preparation of these materials on solid supports[34] (Figure 14). The stage is now set for the Ellman laboratory to create a benzodiazepine library.

Several other approaches to nonpolymeric molecular diversity have recently been published. In pursuit of small-molecule libraries, Nikolaiev *et al.* have used their amino acid encoding format (part 1[1]) with a building block basis set combining both amino acids and other synthetic units to prepare collections of nonpeptidic compounds and peptides refractory to Edman degradation (N-blocked peptides).[35] Representative examples of molecules which have emerged from such non-peptide libraries are shown in Figure 15.

A feature of several of the formats used in the display of synthetic diversity is that the potential ligands are tethered to a solid support. While screening strategies have been developed to exploit this feature, it is frequently desirable to screen compounds in solution. Many groups have engaged in developing releasable linker strategies to solubilize potential ligands. The issue has been addressed by a considerably different strategy by Hobbs DeWitt *et al.*, in which solid-phase chemistry, organic synthesis, and a designed parallel reaction apparatus were utilized for the generation of small-molecule libraries, the individuals of which, were termed "diversomers".[36] Target compounds which included dipeptides, hydantoins, and benzodiazepines were synthesized simultaneously but separately,

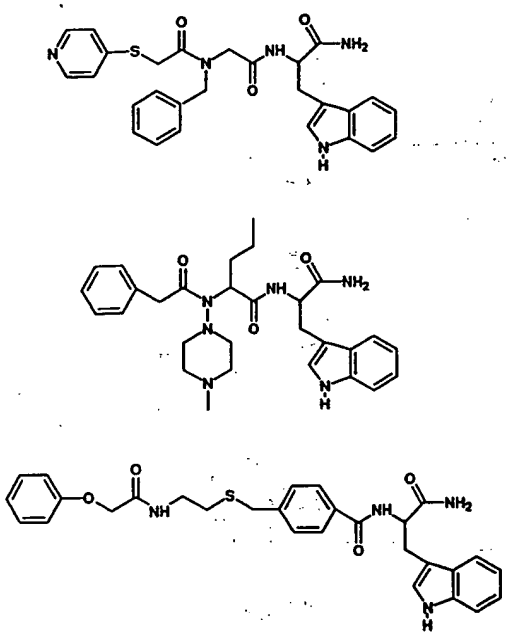


Figure 15. Structure of representative molecules from the Nikolaiev *et al.* nonpeptide library (ref 35).

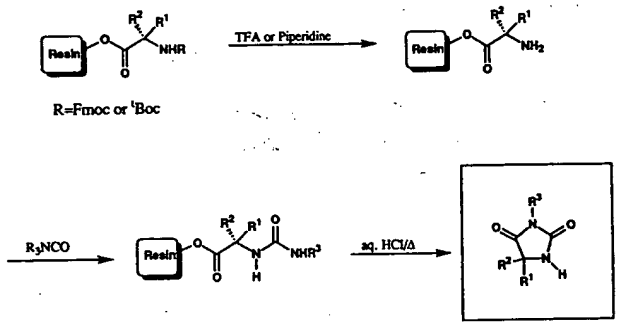


Figure 16. Synthesis of a hydantoin library.

on a solid support in an array format, to generate a collection of up to 40 discrete structurally related compounds. The preparation of hydantoins was carried out as shown in Figure 16. The synthetic strategy is directed through a resin-tethered penultimate product, in which the orchestrated revealing of distal functionality initiates attack on the resin-linking bond to eject the newly formed product into solution. Products which fail to react, should remain attached to the solid phase, and thus aid in product purification. The yields of hydantoins released from the resins in the final step ranged from 4 to 81% on a scale of 0.3–11 mg, which should be sufficient to support most preliminary *in vitro* biological testing. The resulting soluble, small molecules were characterized by traditional means. The authors also note the utility of ^{13}C gel-phase NMR to monitor reaction progress of the resin-bound intermediates.[37,38]

In a similar manner, a general method for multiple, simultaneous synthesis of soluble benzodiazepines was developed (Figure 17). Eight groups of five-amino acid resins were trans-imidated with five groups of eight 2-aminobenzophenone imines to form 40 resin-bound imines. Treatment with TFA liberated 40 discrete benzodiazepines from the resins. The products were obtained in 2–14-mg quantities, corresponding to 9–63% yields with estimated purities of >90%. Though the numbers of compounds involved in the diversomer methodology (~40) are significantly smaller than that which can be prepared by other library strategies (10^4–10^8), this interesting

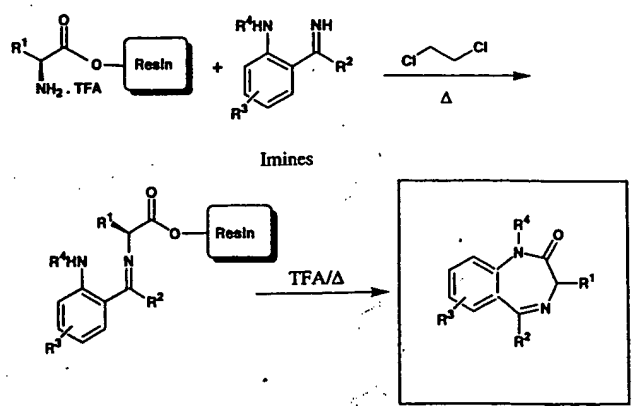


Figure 17. Preparation of a soluble benzodiazepine library.

approach to parallel organic synthesis produces relatively pure materials on a preparative scale in the traditional, soluble format.

Future Innovations

As the field of combinatorial chemistry receives increasing attention from the pharmaceutical establishment, it seems likely that the contents of chemical libraries will continue to evolve to look more and more like the type of compounds which have previously led to drugs. In spite of the complexity which early parts of the process may pose to the combinatorial chemist, a hidden advantage which combinatorially-derived molecules offer is that any "hit" will be readily synthesizable, by definition. This should be contrasted with a natural product driven approach to drug discovery and development, where often the structural complexity of the lead compound hampers the rapid preparation of analog molecules and the acquisition of SAR.

A previous point deserving further emphasis is that the vast universe of synthetic organic reactions are idiosyncratic transformations that fail to afford quantitative yields of unique products. Most synthetic chemistry procedures afford multiple products (regio- and stereoisomers, etc.) in variable yields. If diversity-generating chemistry proceeds ambiguously, how then are the results of small-molecule combinatorial organic syntheses to be understood and appropriate information extracted from library analysis? It may be speculated that encoding techniques will provide one method by which the combinatorial organic chemist can address the practical inefficiencies of chemical synthesis. Instead of envisioning an encoding tag as explicitly specifying the structure of an associated entity, one might consider the tag as *a record of the chemical history of individual library members*. Thus, after encoding the "recipe" or synthetic protocol used in the assembly of a combinatorial library, the library may be screened for active recipes. Once identified as "active", the synthesis would be replicated on a preparative scale, and the product mixture fractionated to identify active product(s). This strategy shifts emphasis from the criterion of singularity in a reaction outcome (a single predictable structure) to reproducibility and compatibility (orthogonality) with chemistry used in the synthesis of the encoding tag and in preparative scaling. The creation of encoded, small-molecule diversity, which can be released from a support (solubilized) while some type of link to the original tag is also maintained, is also likely to be an important area of investigation.

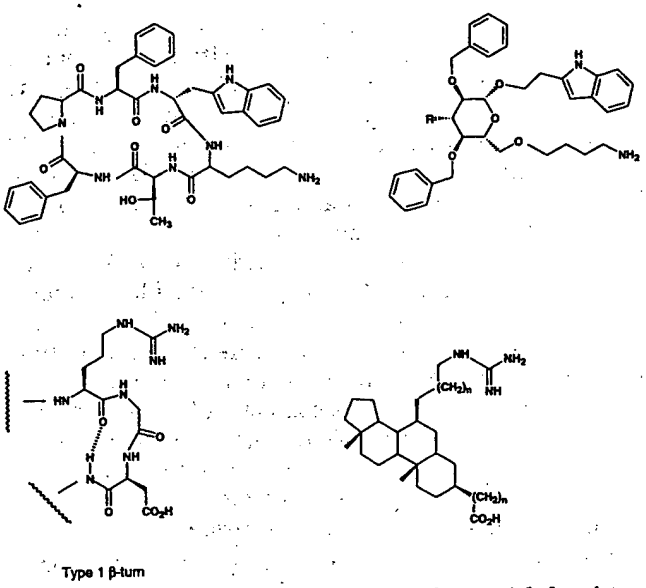


Figure 18. Scaffolds as templates for combinatorial chemistry.

One type of noteworthy chemical strategy which may have a bright future in the combinatorial realm rests on the conceptual extension of the work of Hirschmann, Nicolaou, and Smith and their co-workers into designing nonpeptidal peptidomimetics by the appropriate functionalization of designed scaffoldings (Figure 18). A specific example of this approach is the design and synthesis of a β-D-glucose-based nonpeptide mimetic of a potent cyclic hexapeptide somatostatin agonist.[39,40]

Hirschmann and co-workers have also used a functionalized steroidal template to serve as a backbone for mimicking a type 1 β turn.[41,42] The fact that such a remarkable job of molecular mimicry can be achieved with individual compounds bodes well for the application of this approach to combinatorial methods.

No library will ever be "complete" but instead will sample a subsection of a particular universe of molecular structure and space. In certain situations, libraries may be considered to be starting materials for the construction of new libraries of diversity. It is also useful to consider chemical libraries as collectable or archivable entities. Ideally, one seeks to preserve new compound libraries and use them for a variety of present and future screening needs. As time passes, the combinatorial chemist will be in possession of an accumulating collection of molecular diversity with which to challenge new drug targets. If chemical libraries are to become an item of commerce, a good deal more will have to be learned about their "shelf life" and how best to store them for future use. To date, the shelf life of chemical libraries is an open question.

B. Methods for Screening Combinatorial Libraries

The importance of distinguishing between the two principal applications of combinatorial technologies in ligand discovery, i.e., broad screening versus directed analoging, is particularly relevant to the design of assay methodologies for library evaluation. In searching large, highly diverse libraries for novel lead compounds, a premium is placed on the ability to detect rare ligands that may have modest affinity for the target receptor. The assay strategy may differ in screening analog libraries, since one is trying to develop quantitative SAR on a large number of compounds and to increase the potency of a lead. Regardless of the application, successful use of combinatorial libraries is highly dependent on the sensitivity and specificity of the assays that are used to identify and characterize ligands.

In this section, the various combinatorial library methods will be discussed in terms of the assays that are used. The assay formats are closely matched to the mode of presentation of the diversity. In broad terms, assay procedures can be grouped into three categories: (i) those that rely on affinity purification with an immobilized target receptor, (ii) those in which a soluble receptor binds to tethered ligands, and (iii) those in which soluble compounds are tested for activity, either directly or in competition assays. Each format presents different challenges with regard to the minimum affinity requirements for ligand detection, the demonstration of binding specificity, and the ability to discriminate among compounds in the library on the basis of their affinities for the target.

Isolation of Ligands by Affinity Purification

Recombinant Peptide Libraries. The various systems described in the first part of this series[1] for creating vast libraries of recombinant peptides (commonly referred to as peptide/nucleic acid complexes below) rely on affinity purification to select peptides that bind to a receptor. Two distinct methods have been employed to achieve affinity purification of peptide/nucleic acid complexes. The first involves incubation of a receptor in solution with the mixture of complexes. After allowing sufficient time for binding, the receptor is captured using immobilized streptavidin or an antireceptor antibody.[43,44] The second approach calls for preimmobilization of the receptor on beads, microtiter wells, or a chromatography support, followed by capture of the complexes.[45] In both cases, the use of a solid-support facilitates the separation and washing of receptor-bound complexes.

The method of receptor immobilization is a critical aspect of the successful use of recombinant peptide libraries. Because of the tremendous levels of ligand enrichment attainable through multiple rounds of selection and amplification, peptides that bind to any component of the solid support can be isolated from libraries. Peptides binding to streptavidin,[46,47] antireceptor antibodies,[48] or peptides that exhibit inherent nonspecific binding are readily isolated. Often, screening strategies employing subtractive methods and/or blocking ligands, are necessary to enhance the selection of ligands with desired binding specificity.

To enhance the probability of isolating peptide ligands with biological function, it is important to ensure that the receptor is active (for example, capable of binding its natural ligand) when immobilized. Immobilization of receptor proteins on microtiter wells or beads can be accomplished by passive adherence, covalent attachment, biotinylation and immobilization on streptavidin, or capture with high-affinity nonblocking antibodies. The first three processes often result in inactive proteins. The problem of immobilizing active protein can often be overcome by introducing into the receptor an immobilization handle through genetic engineering techniques. Peptide epitopes for a monoclonal antibody or a motif that allows for site-specific biotinylation of the protein[49] can be fused to proteins for this purpose. Generic immobilization strategies of this type greatly facilitate the creation of a high-density affinity matrix suitable for isolation of ligands.

Some drug-discovery targets may not be readily available as pure soluble receptors. Peptide ligands specific for the integrin IIb/IIIa have been successfully isolated from phage libraries by screening against platelets expressing a high density of this receptor.[50] It remains to be seen whether whole cells bearing receptors or other forms of impure macromolecular target will generally be successfully utilized to identify ligands. It can be anticipated that the problem of isolating non-receptor-binding sequences will be enhanced when dealing with impure forms of receptors.

The choice of using preimmobilized receptor rather than incubation with receptor in solution followed by receptor capture on a solid support may become important in one aspect of library screening. Because the phage and LacI/DNA complex systems are capable of displaying multiple copies of a peptide, multivalent binding can occur during affinity purification if receptors are immobilized at sufficient density. Multivalent binding effectively increases the avidity of the bound complexes and allows the isolation of complexes bearing peptides of lower affinity. Although it is possible that multivalent interactions may occur during the capture step of the two-step procedure, the stoichiometry of the phage or LacI complex, receptor, and the capture reagent need to be carefully controlled. Multivalent binding conditions may be more easily engineered with a high-density preimmobilized receptor.

The multivalency of the phage and LacI systems can be exploited to isolate peptides of modest affinities (K_d values of 1–1000 μM). This feature is particularly important in screening random libraries of peptides greater than six or seven amino acids in length. Libraries that can be routinely made have many fewer members (10^8–10^{10}) than the theoretical number of possible sequences for a given peptide length (for example, there are 1×10^{13} possible 10-mers). In general, peptide ligands for a receptor target consist of families of related sequences with few high affinity ligands in the family. Therefore, when a library is created, it is likely that only lower affinity members of the family will be represented. The ability to identify these low-affinity ligands then permits one to proceed to the secondary phase of discovery, the screening of mutagenesis libraries.

Methods for creating many variants of an initial sequence have been described in part 1. Such libraries will generally contain many low-affinity ligands and perhaps some high-affinity ligands in much lower abundance. A demand is therefore placed on the ability to selectively isolate the highest affinity ligands. Several methods have been described for efficient affinity selection. All are based on the prevention or disruption of multivalent interactions. The use of a low density of immobilized receptor to isolate high-affinity peptide ligands from a phage library containing many low-affinity ligands has been demonstrated.[44] Low receptor density reduces the possibility of multivalent interactions between phage particles and the receptor matrix. The "monovalent-phage" approach[45,51,52] has been successfully employed to isolate high-affinity mutants of human growth hormone displayed on phage. In this approach, phage particles with only a single chimeric pIII protein are created, thereby eliminating multivalent binding to immobilized receptor. Using this approach, mutants with K_d values of less than 5 pM have been identified.

An alternative method of affinity selection that may have advantages over other approaches has been developed.[48,53] Phage (or LacI particles) are allowed to bind in a multivalent fashion to a high density of immobilized receptor. For phage-bearing low-affinity ligands, the peptide on an individual pIII protein may be rapidly dissociating and reassociating, but the phage particle will not dissociate unless all the peptides on pIII are simultaneously in the unbound state. Dissociation of the phage can be initiated by addition of a competing ligand, which prevents rebinding of any individual peptide in the complex. Using a model system with peptides of known affinity, it was demonstrated that phage-bearing high-affinity peptides are retained for a greater length of time than phage with lower affinity sequences.[48] The concentration (and affinity) of the competing ligand, as well as the time and temperature of elution, can be varied to select for ligands of various affinities. This method has the advantage of using a high receptor density to ensure a full sampling of ligands of all affinity classes.

Achieving affinity selection is only part of the process of successfully screening recombinant peptide libraries. After selection, it is necessary to establish the binding specificity and, if possible, the affinity of individual peptides that result from the selection. Various assays have been described, including dot blots,[54] colony lifts,[55] and ELISA's with immobilized phage or immobilized receptor.[44] These methods differ in the minimum ligand affinity that is required for detection. In general, assays in which phage or LacI are immobilized (ELISA's, dot blots, colony lifts) require higher affinity (K_d values < 1 μM) peptides for detecting specific binding, and are therefore useful when such ligands are present in the selected pool. However, for reasons cited above, detection of the specific binding of lower affinity ligands is often necessary. In such cases, assays that use a high density of immobilized receptor are required to allow for multivalent binding and to increase the sensitivity of detection. If high-density receptor matrices are used for affinity purification and assay of individual clones, peptides with K_d values as high as 100–500 μM can be isolated with phage and LacI systems.[48]

An additional assay format has been described for estimating the affinity of peptides displayed by individually selected phage clones.[44] Radiolabeled receptor is first allowed to bind to the phage-borne peptides in solution. A high concentration of competing peptide is then added to prevent further binding, and the dissociation of radiolabeled receptor is followed with time. With a monoclonal antibody model system, a good correlation was observed between dissociation rates and the affinity of the peptide determined by solution methods. It has also been suggested that colony lifts with limited receptor concentration may allow discrimination of individual phage clones on the basis of their peptide's affinity.[55] However, this method may be confounded by differential levels of expression of phage by different colonies.

In summary, the successful identification of ligands from recombinant random peptide libraries depends not only on the nature and size of the libraries but also on effective screening strategies. Selection methods and assays of individual clones vary in their ability to select and detect lower affinity peptides and in the ease with which binding specificity can be determined. High-affinity ligands are most desirable, and initial conditions for screening of random libraries can be adjusted so that only high-affinity ligands are selected. However, for reasons stated above,

engineering selection and assay conditions to allow isolation and detection of lower affinity ligands may be generally a more reliable strategy. These initial peptides can serve as starting points for creating secondary recombinant peptide libraries or as leads for refinement by synthetic chemical combinatorial methods.

Affinity Purification of Mixtures of Soluble Synthetic Compounds. Investigators have employed affinity purification methods to isolate ligands from mixtures of soluble peptides[26,56,57] and oligonucleotides (RNA or DNA).[58-60] In the case of nucleic acid libraries, one takes advantage of the ability to enzymatically amplify the molecules resulting from affinity purification, and as with the recombinant peptide systems, multiple rounds of selection and amplification are used. Theoretical considerations in optimizing conditions for the selection of high-affinity oligonucleotides have been described.[61] The authors illustrate (by way of computer simulation) the importance of nucleic acid and receptor concentrations as well as the efficiency of separating specifically bound molecules. Computer simulations show that, under ideal conditions, rare high-affinity molecules can be isolated from large libraries with relatively few rounds of selection and amplification. There have been a number of examples of successful identification of high-affinity oligonucleotides using this process.[62]

Relatively little work has focused on the affinity purification of ligands from soluble peptide libraries. The isolation of ligands for an anti-gp120 antibody from equimolar mixtures containing 19 or 32 peptides has been reported.[56,57] More recently, the same antibody was used to capture ligands from four mixtures, each of 50 peptides, comprised of unnatural amino acids fused to an encoding L-amino acid peptide strand.[26] After affinity purification, the resulting pool of peptides selected was resolved by HPLC and each peak subjected to Edman sequencing and mass spectrometry analysis. A major limitation of this approach is the sensitivity of these analytical methods. Sufficient peptide (>1–10 pmol) must be recovered in order to determine its sequence, requiring that each library member be present in relatively high amounts in the starting pool and that there be sufficient receptor available to isolate the requisite quantity of each of the high-affinity ligands. In addition, the method requires that selected peptide ligands be resolved chromatographically. While it seems unlikely that this methodology will be extended beyond libraries of modest size (less than a few thousand members), it may prove a useful technique for evaluating secondary (analog) libraries. The proposed approach of creating a library of soluble compounds with attached oligonucleotides tags may allow for the structural identification of minute quantities of compounds isolated by affinity purification.[63,64]

In theory, chromatography of compound mixtures using receptor columns should not only facilitate separation of nonbinding members of the library, but should also allow for the resolution of compounds on the basis of their receptor affinities. Work with various model systems has demonstrated that column retention time can be used as an index of affinity.[65,66] While columns of receptor target have been used in batch affinity purification methods, chromatography to resolve ligands of differing affinities has yet to be applied to screening combinatorial libraries. This method may be better suited to isolation of ligands of moderate affinity.[67] An additional limiting factor in the use of chromatography may be the large amount of receptor required to generate enough theoretical plates to effectively resolve compounds.

Binding of Receptors to Immobilized Ligands

Various methods for creating libraries of compounds attached to solid supports (pins, beads, chips, etc.) have been outlined in part 1.[1] Such libraries are screened by detecting the direct binding of a labeled receptor to an immobilized ligand. The identify of the ligand is either determined directly (by peptide sequencing or mass spectrometry), specified by its spatial location in an array, or deduced by reading an encoding tag.

There are a number of important issues related to solid-phase binding assays with immobilized ligands. First, the ability of a receptor to interact with a tethered ligand may be influenced by the site or nature of its covalent attachment to the solid support. In all of the methods published to date, peptide ligands are attached to a linker and support via the carboxy terminus of the sequence. An obvious example of the limitation imposed by this mode of immobilization would be in screening against the G-protein-linked receptors of various peptide hormones, many of which require a free C-terminal carboxamide for activity. In such a case, it is likely that many peptide analogs that would bind when free in solution would be missed in an assay where the same peptides were immobilized via their C-termini. To circumvent this problem, it is advantageous to have several alternative sites of ligand attachment to the surface. Methods for tethering peptides through their N-termini have been identified.[68] It is likely that the issue of how best to tether molecules to surfaces will become even more important when dealing with libraries of small nonpolymeric organic compounds.

The chemical nature of the linkage between the ligand and support may also affect the receptor-ligand interaction. One needs only to look at the variety of resins that are available for affinity chromatography to appreciate the importance of controlling the receptor–ligand interface. The types of linker groups that have been successfully employed in tethered library assays to date have been noted in part 1 of the series. Whether these linkers will generally provide for optimal presentation of compounds to other receptor systems remains to be seen.

Immobilized ligand assays require that the receptor be labeled in a way that allows for highly sensitive detection of receptor binding. The receptor can either be labeled directly or a secondary labeled reagent with high affinity for the receptor can be used. To date, colorimetric enzymes, radioisotopes, and fluorophores have been used in labeling receptors or secondary reagents. The reagents must be labeled in a way that maintains the activity of the receptor, for instance, its ability to bind a natural ligand. This can be greatly facilitated by creating chimeric recombinant receptors that incorporate peptide epitopes of antibodies or peptide sequences for site-specific radioactive phosphorylation[69] or site-specific biotinylation.[49]

Successful screening of libraries of immobilized synthetic ligands is dependent on the same types of issues as have been previously discussed with respect to evaluating recombinant peptide libraries: i.e., the affinity threshold for detection, the ability to discriminate ligands on the basis of their affinities, and the ability to distinguish specific binding from nonspecific binding. Methods development in this area is in its infancy. In principle, it

should be possible to exploit multivalent binding to detect lower affinity ligands. Multivalent receptors can be created by a number of methods, including genetic fusions to generate bivalent receptor/Fc fusions[70] or through the use of monoclonal antibodies or streptavidin to create cross-linked receptors capable of interaction with more than one immobilized ligand. Optimization of the density of immobilized ligands may be required in order to allow for multivalent binding. As has already been noted, it may be important to be able to isolate relatively low-affinity ligands in the initial screening of random libraries. These compounds can then serve as the basis for further library construction in which the goal is to improve ligand affinity.

Affinity discrimination during the screening of either primary random libraries or secondary (analog) libraries is of obvious importance. There has been little published work on methodology in this area. In principle, low receptor concentrations, competing ligand-mediated dissociation, and/or stringent washing conditions can be utilized to identify the highest affinity ligands. Two issues complicate the use of such methods. The first is the likelihood that each pin, bead, or surface synthesis site does not contain the same amount of compound. With different compound loadings, one must be extremely cautious of using the absolute quantity of bound receptor as an index of a molecule's affinity. As new building-block and coupling chemistries are adapted to combinatorial formats, this may become a more significant problem than it is for high-yielding peptide chemistry. Another complicating feature of the immobilized ligand assay format is the fact that ligands of one particular kind are densely clustered on a surface. Both the association and dissociation rate constants of a receptor/ligand interaction are affected by surface ligand density. The binding of nearby ligands depletes the local receptor concentration and the association kinetics become diffusion limited. After dissociation, receptor rebinding is favored because of the high local-ligand concentration and the apparent dissociation rate is reduced. Theoretical and experimental analyses of these surface binding effects have been undertaken.[71] The impact that these surface binding kinetics will have on the ability to discriminate among library members on pins, beads, or glass surfaces remains to be seen.

The information generated by screening immobilized ligand libraries differs among the various library formats. In the case of bead-based technologies, compounds exceeding a threshold affinity are sampled from a large pool of ligands. Positive information is obtained, i.e., that a particular ligand binds to the receptor. One cannot, however, draw conclusions about the binding affinity of nonselected ligands. The sampling size may not have been large enough to include all high-affinity ligands, or a high-affinity bead may have been missed by the affinity selection method [for example, fluorescence-activated cell sorter (FACS) selection]. By contrast, the multipin and VLSIPS technologies allow one to perform a parallel assay in which data is obtained on every compound that is synthesized. In principle, both positive and negative binding information can be exploited in the design of second-generation compounds.

Incorporation of methods that assess the specificity of binding of ligands is an important aspect of screening random libraries. Screening immobilized ligands by direct receptor binding can lead not only to the identification of ligands of interest (for instance, ligands that compete with the natural ligand) but also to ligands that bind to undesired portions of the receptor or to secondary detection reagents. In the case of libraries of compounds on beads, it may be possible to remove undesired ligands in a subtraction step prior to screening for desired ligands. For compounds on pins or chips, it may be possible to make replicate arrays and test for total binding and nonspecific binding in parallel. Otherwise, sequential assays that first test for receptor binding of any kind, followed by an assessment of nonspecific binding will be required in order to correctly identify compounds that interact with the receptor in a desired manner.

Testing the Activity of Libraries of Soluble Compounds

The classical method of screening for a desired biological activity is to test soluble compounds one at a time in a competition binding assay, enzyme inhibition assay, or in a cell-based bioassay. Such approaches have been applied to library screening by releasing compounds synthesized on pins into microtiter wells, as described in section C of part 1 of this series.[1] A novel application of bead technology has recently been disclosed where compounds on individual beads are released locally onto a lawn of confluent mammalian cells and cause activation of cells in the area surrounding the bead.[72] The bead responsible for cell activation is isolated and a small amount of noncleaved peptide is sequenced to determine its structure. In both of these cases, the principal issue is whether enough compound is released to be detected by the assay. For pins, approximately 100 nmol of peptide can be released into a few hundred microliters of solution, while beads with diameters of $\sim$100 μm can release on the order of 100 pmol of peptide.

Rather than assaying compounds individually, a second approach to screening soluble libraries is to assay compound mixtures. In addition to testing complex pools of soluble peptides (*vide supra*), libraries of oligonucleotides have been successfully screened as soluble mixtures.[73] The most frequently used strategy for screening mixtures of soluble compounds with the goal of ultimately identifying single active molecules is based on the "mimotope" approach, detailed in part 1. The essence of this strategy is that degenerate pools of peptides (or other compounds) are resolved into their most active constituents by an iterative process of testing and resynthesis until a single sequence is identified as having high activity. A variation of the methodology (termed "bogus-coin strategy") has also been described.[74]

There are a number of caveats to using this methodology for testing soluble compound mixtures. In practice, the results of each set of assays do not typically indicate a preference for a unique residue at any position within the sequence. Rather, comparable assay results may be obtained for several different amino acid substitutions and some decision must be made as to which of these partial solutions should be fully resolved. The number of peptide mixtures to be synthesized and tested in this protocol expands dramatically as the number of alternative sequences selected for complete resolution at each cycle is increased. Moreover, the deconvolution of different partial solutions may frequently produce divergent resolved sequences, in part because the contribution of each amino acid to the peptide–receptor interaction is typically

dependent on other non-neighboring residues within the ligand. The problem of identifying the most potent ligand in a complex mixture by this type of iterative pathway is exacerbated by the relative abundance of lower affinity ligands that represent local binding optima.

Originally designed for identifying antibody ligands, the mimotope strategy has primarily been used for libraries of six to eight building blocks in length. It is not clear that ligands of this size will be optimal for other types of receptors (although success with opioid receptors[75] and other targets have been reported). As the length of the compounds in the library increases, resynthesis and testing of pools becomes more cumbersome.

Perhaps the greatest limitation of this methodology is the fact that the activity of a given pool is based on the cumulative activity of all the compounds in the pool; i.e., pools with the same activity may contain many low-affinity compounds or a few high-affinity compounds. For this reason, the methodology is greatly facilitated if the minimal fragment having activity is comprised of the same number of building blocks as used in constructing each library member (e.g., a uniquely active tetramer is more easily resolved from a tetrapeptide library than a hexamer library). Alternatively, the identification of active peptide(s) is facilitated if the receptor has specific requirements for a fixed position within a peptide ligand (e.g., the N or C termini). If neither of these conditions is true, it may be necessary to test many or all of the possible initial pools with two adjacent or nonadjacent fixed residues. This drastically increases the number of initial pools that need to be synthesized but increases the probability that a critical residue(s) is fixed in at least one pool to allow that pool to differentiate itself. It must be kept in mind that any pool identified as having the greatest activity may be composed of many moderately active compounds and that the most active compound(s) may reside in other pools.

Testing of mixtures of soluble compounds is also limited by the concentration of individual test compounds that can be achieved in the initial pools. Pools containing as many as 160 000 different peptides have been tested with each member being present at ~10 nanomolar concentration.[76] Because of limitations on the solubility of the total pool, the concentrations of individual compounds present in increasingly larger libraries must be correspondingly diminished. This will ultimately limit the ability to identify the activity of compounds with modest potencies.

While the current methods for testing mixtures of soluble compounds have certain drawbacks, screening soluble libraries does have the decided advantage of avoiding the problems associated with assaying tethered molecules in other combinatorial technologies. Conventional binding and enzyme and cell-based assays (including those with poorly defined biochemical targets) can be used to test the activity of soluble compounds. It is likely that in the future, encoding strategies will be employed to allow more facile screening of soluble molecules. In the simplest format, single encoded beads can be dispensed into microtiter wells. The compounds can then be released from the beads and tested for activity, with the identity of the most active compound(s) being deduced by decoding the tag attached to the bead(s). To test large libraries of soluble compounds, mixed pools of encoded beads can be created. At each round of testing, only a fraction of the compound is cleaved from each bead. Active pools of beads are pursued by further subdividing the beads, partially releasing the compounds, and testing for activity. In the end, a single bead is identified with the greatest activity and the encoding tag is then read.[77]

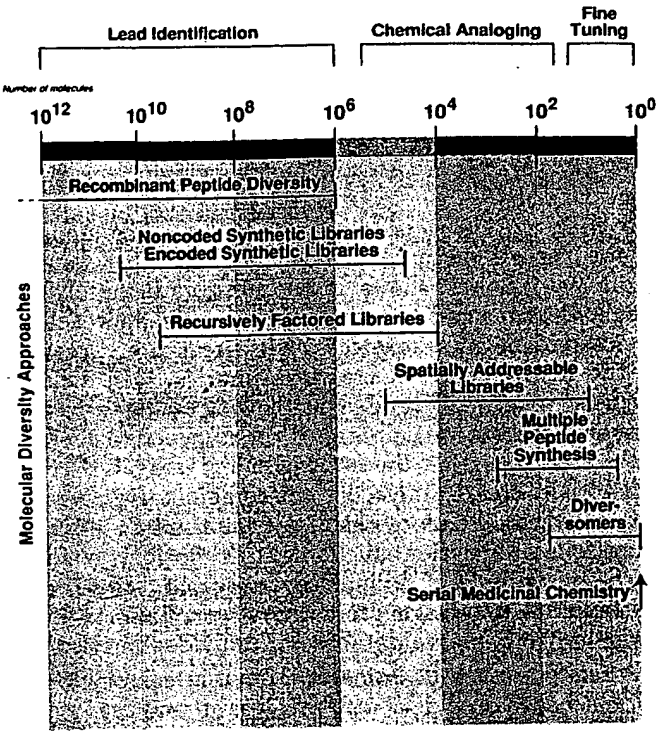


Figure 19.

C. Integration of Combinatorial Technologies for Drug Discovery and Future Directions of the Field

Regardless of whether the objective is a broad discovery search or analoging a known lead, a key aspect in the successful application of combinatorial technologies to drug discovery is the requirement for having a closely linked, coordinated process for the integration of synthesis and screening. The creation and evaluation of molecular diversity are two sides of the same coin. It is still premature to speculate on which type of format will prove most suitable for a particular type of discovery/development problem. Over time, a cataloging of successes and failures will clarify this issue. In all likelihood, command of a collection of combinatorial tools will be required for general success. One may imagine a "spectrum of molecular diversity" stretching from few to many molecules (see Figure 19). Serial synthesis (contemporary medicinal chemistry) operates on a few molecules (far right of chart). We expect that each individual combinatorial tool/format will turn out to be most valuable within proscribed regions of numbers of molecules. Hence, recombinant peptide diversity is particularly suitable for generating and screening large libraries of $>10^8$ compounds. The VLSIPS chip technology, although capable of displaying vast arrays, is primarily an analoging tool and is most useful for evaluating 10^2–10^4 compounds. The parallel organic synthesis methods of Hobbs DeWitt *et al.*[36] are applicable for tens of compounds. Encoded synthetic libraries appear to be relevant in broad diversity searching and may also prove useful in narrower optimization strategies. Given the repetitive nature of many of the manipulations required for library construction, an on-going priority will be to address the possibility of automating as many aspects of the generation/evaluation process as possible. Growth in

library size also impacts directly on the physical size of compound collections and on the amounts of target receptor required for screening. These pressures will continue to drive the field toward minaturization and exploitation of nanotechnologies.

The power of combinatorial technologies in generating huge numbers of compounds suggests that in a lead-discovery mode, less preconceived bias need be brought to the process of making molecules. Another way of expressing this is as follows: let the numbers do the talking. Due to the time and effort required in serial approaches, each target molecule is selected with great care. Because of the relative ease in creating libraries, little risk is incurred, or effort expended, in allowing a wide variety of building blocks to participate in diversity generation. Since there is less up-front investment in any individual combinatorially created molecule, the combinatorial chemist can afford to take more risks. We can think in terms of a portfolio of libraries which might be routinely applied to the initiation of a drug-discovery search. This is not dissimilar to selecting a preliminary screening sample of diversity from a large database of individual molecules.

A related, but still immature, issue in combinatorial approaches to drug discovery revolves around the idea of "quantitation of diversity". An understanding of the concept of "measuring" molecular diversity could impact on designing libraries to contain maximal structural diversity. This notion has arisen previously in deciding which few representative, highly diverse compounds to select out of large database collections, when setting up groups of preliminary screening samples. The huge numbers involved in combinatorial approaches intensifies this issue. A number of interesting approaches to the diversity quantitation problem can be expected to emerge.

One working drug-discovery paradigm might be based on initially employing a portfolio of biological diversity (peptide libraries) together with standard chemical libraries (various-sized cyclic peptides, cysteinyl-linked cyclics, etc.), peptides with carboxyl or carboxamide display, synthetic polymeric diversity, as well as large libraries of semirigid and acyclic small molecules prepared by COS. Over time, favored libraries and directions would be expected to emerge. As the sophistication of combinatorial organic synthesis grows, the origins of a molecular structure as either "combinatorially or serially derived" will gradually become transparent.

Another area where considerable effort must be applied is in the registry of libraries and individual library members. It is unclear that library compounds should be registered and documented for testing in the same ways as serially produced compounds historically have been, but exactly what changes are necessary remain to be determined. Vast numbers of compounds have been and are being created; keeping track of these and their corresponding biological activities will require innovative database-management techniques. Additionally, nomenclature needs to be developed by which one can simply express the constitution, scope, and nature of chemical libraries. Legal issues, including the patenting and documentating of libraries and their component members, will need to be pioneered.

As repeatedly emphasized, it is obvious but imperative to have efficient means of evaluating the molecular diversity which is generated. Different assay techniques will be format specific. Assays must clearly discriminate specific from nonspecific binding. Since in a broad screening mode, one is almost always sampling a small percentage of the entire universe of diversity (10^{10} peptides are only 0.1% of the universe of 10-mers), it is crucial that appropriate assay techniques be competent to detect modest affinity ligands. The identification of weak binders in any of the aforementioned approaches is very important and should lead directly to preparation of secondary libraries in which original "hits" will become the centerpiece for more focused diversity creation. This is a consequential issue, since application of combinatorial technologies are best viewed as an iterative process and not a singular event. As the emphasis shifts to analog evaluation, assays must be capable of affinity discrimination between closely related library members. The tools of molecular biology have permitted the molecular engineering of targets to serve the purposes of screening. The rapid introduction of targets into a screening mode will require generic techniques for their handling, and manipulation of molecular targets by genetic engineering will continue to play a crucial role in marrying library evaluation and synthesis. Though combinatorial technologies may soon prove their worth in the drug-discovery process by delivering new leads quickly and cheaply, in order to completely fulfill the promise of "making drugs", an important question will be whether some of the common major obstacles to drug development (e.g., cell penetration, bioavailability, pharmacokinetics, metabolism) can be productively addressed through the application of combinatorial approaches (i.e., *combinatorial drug development*).

In the coming years, cloning and sequencing of the human genome promises that an unprecedented abundance of newly discovered proteins will become available as potential drug targets. Gaining even more prominence than it now assumes will be the issue of discriminating among a myriad of receptors and enzymes to identify valid targets for drug discovery. The ability to access potent and specific ligands for these targets will guide this process by untangling the physiological relevance of endogenous biochemical pathways. Combinatorial methods will be called upon to provide such molecules to quickly and cheaply drive target validation. In this manner, the identification of leads will benefit from a significant, but hidden, benefit which emerges from combinatorial screening; hits derived from chemical libraries should be readily amenable to combinatorial analoging.

Certain drug targets may present more or less of a historical precedent with respect to the likelihood of successfully identifying a tight binding ligand through the use of known pharmacophores. For example, the search for specific enzyme inhibitors may be facilitated by the intentional enrichment in the combinatorial synthetic process of building blocks containing known inhibitory pharmacophores. Particularly important or common types of drug targets may justify having on hand special libraries which are somewhat specific (i.e., a peptidyl hydroxyethylamine library for aspartyl- and metalloprotease inhibition[78-80]). On the other hand, in areas where there is less current information (e.g., antagonism of protein–protein or carbohydrate–protein interactions), a wider scope of diversity search should be taken until consistent patterns begin to emerge. In the case of newer, less explored target groups, combinatorial technologies can be expected to assist in unearthing new pharmacophore

classes and to help establish an understanding of drug design for new types of targets.

Combinatorial technologies diverge sharply from historical precedent through a change in emphasis from the consideration of individual molecules to thinking in terms of populations of molecules. A common, but false, intuitive believe is that combinatorial chemistry is necessarily a random, screening search; the antithesis of rational drug design. In fact, all libraries are biased in some ways. All drug company compound files are biased by the historical programs of that institution, since a disproportionate share of compounds of particular types will have been deposited. *The notion of intentionally biasing a chemical library is a form of drug design*, but again not applied to individuals but rather to groups or populations of molecules. If a scientist hypothesizes on the basis of structural information that the current lead molecule contains a type II β-turn motif, then rather than performing two or three serial tests of this idea, the combinatorial chemist might create a library of narrow diversity utilizing a basis set of β-turn mimetics and thus interrogate many slightly different regions of conformational space simultaneously. The drug design of populations versus individuals is analogous to fishing with a net rather than just a hook. As more knowledge of workable strategies for combinatorial synthesis are understood, it is expected that structural and computational input and other rational design information will be integrated into a broad combinatorial medicinal chemistry approach.

Gaining a full appreciation of the issues and difficulties which must be surmounted in order to perform useful combinatorial organic synthesis will initially be a relatively slow process, especially because the important strategies and decision points differ so markedly from traditional organic synthesis. Retrocombinatorial analysis of existing pharmacophores and other important structures should assist in decision making; both in choosing routes of forward synthesis and in synthetic target selection. If combinatorial techniques are indeed to become a useful shortcut to new leads and optimized compounds, then one key implied goal of combinatorial organic synthesis is to intersect the pathway of modern medicinal chemistry upon which compounds move from the early discovery stages to clinical candidacy. Rich incentives await those who are able to mass produce important biologically active molecules quickly and cheaply. Not surprisingly, an aggressive, worldwide effort to understand and master this field has already begun.

This Perspective has been restricted to a consideration of the impact of combinatorial technologies on medicinal chemistry/drug discovery and development. From the point of view of applicability of the technologies, this is an artificially narrow view. Combinatorial processes will become important in diagnostic medicine,[81] agricultural chemistry, food chemistry, immunology, molecular biology, polymer studies, inorganic synthesis, and many other fields. Though the field of "combinatorial chemistry" is chronologically a new enterprise, the evolution of thought in this fertile area continues to outrace the experimental reduction to practice of many ideas. One may reasonably ask "why are combinatorial technologies happening now?". The answer is probably complex and beyond the scope of this Perspective. Nevertheless, the explosive recent interest in the application of combinatorial technologies to drug discovery is symptomatic of an idea whose time has come. Because the issues which confront the medicinal chemist differ so radically from historical approaches, the combinatorial field will no doubt continue to provide impetus and stimulation for the formulation of new concepts and ideas.

Acknowledgment. We thank Drs. A. Zaffaroni and G. Ringold for their interest, insights, and support in undertaking the application of combinatorial technologies to drug discovery and Mrs. Terri Nebozuk for help in preparing this manuscript.

References

(1) Part 1: Gallop, M. A.; Barrett, R. W.; Dower, W. J.; Fodor, S. P. A.; Gordon, E. M. Applications of Combinatorial Technologies to Drug Discovery. 1. Background and Peptide Combinatorial Libraries. *J. Med. Chem.* **1994**, *37*, 1233–1251.

(2) Barker, P. L.; Bullens, S.; Bunting, S.; Burdick, D. J.; Chan, K. S.; Deisher, T.; Eigenbrot, C.; Gadek, T. R.; Lipari, M. T.; *et al.* Cyclic RGD Peptide Analogues as Antiplatlet Antithrombotics. *J. Med. Chem.* **1992**, *35*, 2040–2048.

(3) Al-Obeidi, F.; De L Castrucci, A. M.; Hadley, M. E.; Hruby, V. J. Potent and Prolonged Acting Cyclic Lactam Analogues of an α-Melanotropin: Design Based on Molecular Dynamics. *J. Med. Chem.* **1989**, *32*, 2555–2561.

(4) Wiley, R. A.; Rich, D. H. Peptidomimetics Derived from Natural Products. *Med. Res. Rev.* **1993**, *13*, 327–384.

(5) Giannis, A.; Kolter, T. Peptidomimetics for Receptor Ligands—Discovery, Development, and Medicinal Perspectives. *Angew. Chem., Int. Ed. Engl.* **1993**, *32*, 1244–1267.

(6) Kates, S. A.; Sole, N. A.; Johnson, C. R.; Hudson, D.; Barany, G.; Albericio, F. A Novel, Convenient, Three-Dimensional Orthogonal Strategy for Solid Phase Synthesis of Cyclic Peptides. *Tetrahedron Lett.* **1993**, *34*, 1549–1552.

(7) Polinsky, A.; Cooney, M. G.; Toy-Palmer, A.; Osapay, G.; Goodman, M. Synthesis and Conformational Properties of the Lanthionine Bridged Opioid Peptide [D-Ala_L^2,Ala_L^5]Enkephalin as determined by NMR and Computer Simulations. *J. Med. Chem.* **1992**, *35*, 4185–4194.

(8) Albericio, F.; Hammer, R. P.; Garcia-Echeverria, C.; Molins, M. A.; Chang, J. L.; Munson, M. C.; Pons, M.; Giralt, E.; Barany, G. Cyclization of Disulfide-Containing Peptides in Solid Phase Synthesis. *Int. J. Pept. Protein Res.* **1991**, *37*, 402–413.

(9) Marlowe, C. K. Peptide Cyclization on TFA Lable Resin Using the Trimethylsilyl (TMSE) Ester as an Orthogonal Protecting Group. *Bioorg. Med. Chem. Lett.* **1993**, *3*, 437–440.

(10) Trzeciak, A.; Bannwarth, W. Synthesis of 'Head to Tail' Cyclized Peptides on Solid Support by FMOC Chemistry. *Tetrahedron Lett.* **1992**, *33*, 4557–4560.

(11) Andrews, P. C.; Boyd, J.; Ogorzalek Loo, R.; Zhao, R.; Zhu, C.-Q.; Grant, K.; Williams, S. Synthesis of Uniform Peptide Libraries and Methods for Physico-Chemical Analysis. Poster presented at the Meeting of the Protein Society, San Diego, July 1993.

(12) Metzger, J. W.; Wiesmuller, K.-H.; Gnau, V.; Brunjes, J.; Jung, G.; Ion-Spray Mass Spectrometry and High Performance Liquid Chromatography–Mass Spectrometry of Synthetic Peptide Libraries. *Angew. Chem., Int. Ed. Engl.* **1993**, *32*, 894–896.

(13) Zuckermann, R. N.; Kerr, J. M.; Siani, M. A.; Banville, S. C. Design, Construction and Application of a Fully Automated Equimolar Peptide Mixture Synthesizer. *Int. J. Pept. Protein Res.* **1992**, *40*, 497–506.

(14) Fodor, S. P. A.; Read, J. L.; Pirrung, M. C.; Stryer, L.; Lu, A. T.; Solas, D. Light-Directed, Spatially Addressable Parallel Chemical Synthesis. *Science* **1991**, *251*, 767–773.

(15) Zuckermann, R. N.; Kerr, J. M.; Siani, M. A.; Banville, S. C.; Santi, D. V. Identification of Highest-Affinity Ligands by Affinity Selection from Equimolar Peptide Mixtures Generated by Robotic Synthesis. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 4505–4509.

(16) Freidinger, R. M.; Perlow, D. S.; Veber, D. F. Protected Lactam-Bridged Dipeptides for Use as Conformational Constraints in Peptides. *J. Org. Chem.* **1982**, *47*, 104–109.

(17) Crowley, J. I.; Rapoport, H. Solid-Phase Organic Synthesis: Novelty or Fundamental Concept. *Acc. Chem. Res.* **1976**, *9*, 135–144.

(18) Leznoff, C. C. The Use of Insoluble Polymer Supports in General Organic Synthesis. *Acc. Chem. Res.* **1978**, *11*, 327.

(19) Frechet, J. M. J. Synthesis and Applications of Organic Polymers as Supports and Protecting Groups. *Tetrahedron* **1981**, *37*, 663–683.

(20) Hodge, P. Organic Reactions Using Polymer-Supported Catalysts, Reagents or Substrates. In *Syntheses and Separations Using Functional Polymers*; Sherrington, D. C., Hodge, P., Eds.; John Wiley and Sons: New York, 1988, pp 43–122.

(21) Cho, C. Y.; Moran, E. J.; Cherry, S. R.; Stephans, J. C.; Fodor, S. P. A.; Adams, C. L.; Sundaram, A.; Jacobs, J. W.; Schultz, P. G. An Unnatural Biopolymer. *Science* **1993**, *261*, 1303–5.

(22) Simon, R. J.; Kania, R. S.; Zuckermann, R. N.; Huebner, V. D.; Jewell, D. A.; Banville, S.; Ng, S.; Wang, L.; Rosenberg, S.; Marlowe, C. K.; Spellmeyer, D. C.; Tan, R.; Frankel, A. D.; Santi, D. V.; Cohen, F. E.; Bartlett, P. A. Peptoids: A Modular Approach to Drug Discovery. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 9367–9371.
(23) Bartlett, P. A.; Santi, D. V.; Simon, R. J. Int. Patent WO91/19735, 1991.
(24) Zuckermann, R. N.; Kerr, J. M.; Kent, S. B. H.; Moos, W. H. Efficient method for the preparation of peptoids[oligo(N-substituted glycines)] by submonomer solid-phase synthesis. *J. Am. Chem. Soc.* **1992**, *114*, 10646–10647.
(25) Zuckermann, R. N. Presented at the Third International Symposium on Solid Phase Synthesis and Complementary Technologies, Oxford, U.K., 31 August–4 September, 1993.
(26) Kerr, J. M.; Banville, S. C.; Zuckermann, R. N. Encoded Combinatorial Peptide Libraries Containing Non-Natural Amino Acids. *J. Am. Chem. Soc.* **1993**, *115*, 2529–2531.
(27) Rich, D. H. Peptidase Inhibitors. In *Comprehensive Medicinal Chemistry*; Hansch, C., Sammes, P. G., Taylor, J. B., Eds.; Pergamon Press: Oxford, 1990, Vol. 2.
(28) Bartlett, P. A.; Marlowe, C. K. Possible Role for Water Dissociation in the Slow Binding of Phosphorous-Containing Transition-State-Analogue Inhibitors of Thermolysin. *Biochemistry* **1987**, *26*, 8553–8561.
(29) Morgan, B. P.; Scholtz, J. M.; Ballinger, M. D.; Zipkin, I. D.; Bartlett, P. A. Differential Binding Energy: A Detailed Evaluation of the Influence of Hydrogen-Bonding and Hydrophobic Groups on the Inhibition of Thermolysin by Phosphorous-Containing Inhibitors. *J. Am. Chem. Soc.* **1991**, *113*, 297–307.
(30) Campbell, D. A.; Bermak, J. C. Phosphonate Ester Synthesis Using a Modified Mitsunobu Condensation. *J. Org. Chem.* **1994**, *59*, 658–660. Campbell, D. A. The Synthesis of Phosphonate Esters, an Extension of the Mitsunobu Reaction. *J. Org. Chem.* **1992**, *57*, 6331–6335.
(31) Hagihara, M.; Anthony, N. J.; Stout, T. J.; Clardy, J.; Schreiber, S. L. Vinylogous Polypeptides: An Alternative Peptide Backbone. *J. Am. Chem. Soc.* **1992**, *114*, 6568–6570.
(32) Smith, A. B., III; Keenan, T. P.; Holcomb, R. C.; Sprengler, P. A.; Guzman, M. C.; Wood, J. L.; Carroll, P. J.; Hirschmann, R. Design, Synthesis and Crystal Structure of a Pyrrolinone-Based Peptidomimetic Possessing the Conformation of a β-Strand: Potential Applications to the Design of Novel Inhibitors of Proteolytic Enzymes. *J. Am. Chem. Soc.* **1992**, *114*, 10672–10674.
(33) Bunin, B. A.; Ellman, J. A. A General and Expedient Method for the Solid-Phase Synthesis of 1,4-Benzodiazepine Derivatives. *J. Am. Chem. Soc.* **1992**, *114*, 10997–10998.
(34) Ellman, J. A. Pharmaceutical Manufacturers Association, Drug Discovery Management Subsection, September 19–21, 1993, Philadelphia, PA.
(35) Nikolaiev, V.; Stierandova, A.; Krchnak, V.; Seligmann, B.; Lam, K. S.; Salmon, S. E.; Lebl, M. Peptide-Encoding for Structure Determination of Nonsequenceable Polymers Within Libraries Synthesized and Tested on Solid-Phase Supports. *Pept. Res.* **1993**, *6*, 161–170.
(36) Hobbs DeWitt, S.; Kiely, J. S.; Stankovic, C. J.; Schroeder, M. C.; Reynolds Cody, D. M.; Pavia, M. R. "Diversomers": An Approach to Nonpeptide, Nonoligomeric Chemical Diversity. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 6909–6913.
(37) Mapelli, C.; Swerdloff, M. D. Monitoring of Conformational and Reaction Events in Resin-Bound Peptides by Carbon-13 NMR Spectroscopy in Various Solvents. In *Peptides. Proceedings of the 21st European Peptide Symposium*; Giralt, E. and Andreu, D., Eds., ESCOM, Leiden, The Netherlands, 1990; pp 316–319.
(38) Giralt, E.; Rizo, J.; Pedroso, E. Application of Gel-Phase ^{13}C-NMR to Monitor Solid Phase Peptide Synthesis. *Tetrahedron* **1984**, *40*, 4141–4152.
(39) Nicolaou, K. C.; Salvino, J. M.; Raynor, K.; Pietranico, S.; Reisine, T.; Freidinger, R. M.; Hirschmann, R. Design and Synthesis of a Peptidomimetic Employing β-D-Glucose for Scaffolding. *Peptides* **1989**, 881–884.
(40) Hirschmann, R.; Nicolaou, K. C.; Pietranico, S.; Salvino, J.; Leahy, E. M.; Sprengeler, P. A.; Furst, G.; Smith, A. B.; III; Strader, C. D.; Cascieri, M. A.; Candelore, M. R.; Donaldson, C.; Vale, W.; Maechler, L. Nonpeptidal Peptidomimetics with a β-D-Glucose Scaffolding. A Partial Somatostatin Agonist Bearing a Close Structural Relationship to a Potent, Selective Substance P Antagonist. *J. Am. Chem. Soc.* **1992**, *114*, 9217–9218.
(41) Hirschmann, R.; Sprengeler, P. A.; Kawasaki, T.; Leahy, J. W.; Shakespeare, W. C.; Smith, A. B., III, The First Design and Synthesis of a Steroidal Peptidomimetic. The Potential Value of Peptidomimetics in Elucidating the Bioactive Conformation of Peptide Ligands. *J. Am. Chem. Soc.* **1992**, *114*, 9699–9701.
(42) Hirschmann, R.; Sprengeler, P. A.; Kawasaki, T.; Leahy, J. W.; Shakespeare, W. C.; Smith, A. B., III, The Versatile Steroid Nucleus: Design and Synthesis of a Peptidomimetic Employing this Novel Scaffold. *Tetrahedron* **1993**, *49*, 3665–3676.
(43) Parmley, S. F.; Smith, G. P. Antibody-Selectable Filamentous Fd Phage Vectors: Affinity Purification of Target Genes. *Gene* **1988**, *73*, 305–318.
(44) Barrett, R. W.; Cwirla, S. E.; Ackerman, M. S.; Olson, A. M.; Peters, E. A.; Dower, W. J. Selective Enrichment and Characterization of High Affinity Ligands from Collections of Random Peptides on Filamentous Phage. *Anal. Biochem.* **1992**, *204*, 357–364.
(45) Bass, S.; Green, R.; Wells, J. A. Hormone Phage: An Enrichment Method for Variant Proteins with Altered Binding Properties. *Proteins: Struct. Funct. Genet.* **1990**, *8*, 309–314.
(46) Devlin, J. J.; Panganiban, L. C.; Devlin, P. E. Random Peptide Libraries: A Source of Specific Protein Binding Molecules. *Science* **1990**, *249*, 404–406.
(47) Kay, B. K.; Adey, N. B.; He, Y.-S.; Manfredi, J. P.; Mataragnon, A. H.; Fowlkes, D. M. An M13 Phage Library Displaying Random 38-Amino-Acid Peptides as a Source of Novel Sequences with Affinity to Selected Targets. *Gene* **1993**, *128*, 59–65.
(48) Barrett, R. W.; Dower, W. J. Unpublished.
(49) Schatz, P. J. Use of peptide libraries to map the substrate specificity of a peptide-modifying enzyme: A 13 residue consensus peptide specifies biotinylation in *Escherichia coli*. *Bio/Tech.* **1993**, *11*, 1138–1143.
(50) Doyle, M. Chiron Corporation. Presented at San Francisco Bay Area Peptide Society Meeting, December 1992.
(51) Lowman, H. B.; Bass, S. H.; Simpson, N.; Wells, J. A. Selecting high-affinity binding proteins by monovalent phage display. *Biochemistry* **1991**, *30*, 10832–10838.
(52) Garrard, L. J.; Yang, M.; O'Connell, M. P.; Kelley, R. F.; Henner, D. J. Fab Assembly and Enrichment in a Monovalent Phage Display System. *Bio/Technology* **1991**, *9*, 1373–1377.
(53) Hawkins, R. E.; Russell, S. J.; Winter, G. Selection of Phage Antibodies by Binding Affinity: Mimicking Affinity Maturation. *J. Mol. Biol.* **1992**, *226*, 889–896.
(54) Kang, A. S.; Barbas, C. F.; Janda, K. K.; Benkovic, S. J.; Lerner, R. A. Linkage of Recognition and Replication Functions by Assembling Combinatorial Antibody Fab Libraries Along Phage Surfaces. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 4363–4366.
(55) Christian, R. B.; Zuckermann, R. N.; Kerr, J. M.; Wang, L.; Malcolm, B. A. Simplified Methods for Construction, Assessment and Rapid Screening of Peptide Libraries in Bacteriophage. *J. Mol. Biol.* **1992**, *227*, 711–718.
(56) Zuckermann, R. N.; Kerr, J. M.; Siani, M. A.; Banville, S. C.; Santi, D. V. Identification of Highest-Affinity Ligands by Affinity Selection from Equimolar Peptide Mixtures Generated by Robotic Synthesis. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 4505–4509.
(57) Kerr, J. M.; Banville, S. C.; Zuckermann, R. N. Identification of Antibody Mimotopes Containing Non-Natural Amino Acids by Recombinant and Synthetic Peptide Library Affinity Selection Methods. *Bioorg. Med. Chem. Lett.* **1993**, *3*, 463–468.
(58) Tuerk, C.; Gold, L. Systematic Evolution of Ligands by Exponential Enrichment: RNA Ligands to Bacteriophage T4 DNA Polymerase. *Science* **1990**, *249*, 505–510.
(59) Ellington, A. D.; Szostak, J. W. *In vitro* Selection of RNA molecules That Bind Specific Ligands. *Nature* **1990**, *346*, 818–822.
(60) Sherman, M. I.; Bertelsen, A. H.; Cook, A. F. Protein Epitope Targeting: Oligonucleotide Diversity and Drug Discovery. *Bioorg. Med. Chem. Lett.* **1993**, *3*, 469–475.
(61) Irvine, D.; Tuerk, C.; Gold, L. Selexion: Systematic Evolution of Ligands by Exponential Enrichment with Integrated Optimization by Non-Linear Analysis. *J. Mol. Biol.* **1991**, *222*, 739–761.
(62) Tuerk, C.; MacDougal, S.; Gold, L. RNA Pseudoknots that Inhibit Human Immunodeficiency Virus Type 1 Reverse Transcriptase. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 6988–6992.
(63) Brenner, S.; Lerner, R. A. Encoded Combinatorial Chemistry. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 5181–5183.
(64) Nielsen, J.; Brenner, S.; Janda, K. D. Synthetic Methods for the Implementation of Encoded Combinatorial Chemistry. *J. Am. Chem. Soc.* **1993**, *115*, 9812–9813.
(65) Arnold, F. H.; Schofield, S. A.; Blanch, H. W. Analytical Affinity Chromatography: I. Local Equilibrium Theory and the Measurement of Association and Inhibition Constants. *J. Chromatogr.* **1986**, *355*, 1–12.
(66) Arnold, F. H.; Blanch, H. W. Analytical Affinity Chromatography: II. Rate Theory and the Measurement of Biological Binding Kinetics. *J. Chromatogr.* **1986**, *355*, 13–27.
(67) Zopf, D.; Ohlson, S. Weak-Affinity Chromatography. *Nature* **1990**, *346*, 87–88.
(68) Holmes, C. P. Polymer Reversal on Solid Surfaces. U.S. Pat. 5,242,974.
(69) Li, B.-L.; Langer, J. A.; Schwartz, B.; Petska, S. Creation of Phosphorylation Sites in Proteins: Construction of a Phosphorylatable Human Interferon α. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 558–562.
(70) Hollenbaugh, D.; Chalupny, N. J.; Aruffo, A. Recombinant Globulins: Novel Research Tools and Possible Pharmaceuticals. *Curr. Opinion Immunol.* **1992**, *4*, 216–219.
(71) Erickson, J.; Goldstein, B.; Holowka, D.; Baird, B. The Effect of Receptor Density on the Forward Rate Constant for Binding of Ligands to Cell Surface Receptors. *Biophys. J.* **1987**, *52*, 657–662.
(72) Jayawickreme, C. K.; Graminski, G. F.; Quillan, J. M.; Lerner, M. R. Creation and Functional Screening of a Multi-Use Peptide Library. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 1614–1618.

(73) Ecker, D. J.; Vickers, T. A.; Hanecak, R.; Driver, V.; Anderson, K. Rational Screening of Oligonucleotide Combinatorial Libraries for Drug Discovery. *Nucleic Acids Res.* 1993, *21*, 1853–1856.
(74) Blake, J.; Litzi-Davis, L. Evaluation of Peptide Libraries: An Iterative Strategy to Analyze the Reactivity of Peptide Mixture with Antibodies. *Bioconj. Chem.* 1992, *3*, 510–513.
(75) Dooley, C. T.; Chung, N. N.; Schiller, P. W.; Houghten, R. A. Acetalins: Opioid Receptor Antagonists Determined Through the Use of Synthetic Peptide Combinatorial Libraries. *Proc. Natl. Acad. Sci. U.S.A.* 1993, *90*, 10811–10815.
(76) Houghten, R. A.; Appel, J. R.; Blondelle, S. E.; Cuervo, J. H.; Dooley, C. T.; Pinilla, C. The Use of Synthetic Peptide Combinatorial Libraries for the Identification of Bioactive Peptides. *BioTechniques* 1992, *13*, 412–421.
(77) Salmon, S. E.; Lam, K. S.; Lebl, M.; Kandola, A.; Khattri, P. S.; Wade, S.; Patek, M.; Kocis, P.; Krchnak, V.; Thorpe, D.; Felder, S. Discovery of Biologically Active Peptides in Random Libraries: Solution-Phase Testing After Staged Orthogonal Release from Resin Beads. *Proc. Natl. Acad. Sci. U.S.A.* 1993, *90*, 11708–11712.
(78) Gordon, E. M.; Godfrey, J. D.; Pluscec, J.; Von Langen, D.; Natarajan, S. Design of Peptide Derived Amino Alcohols as Transition-State Analog Inhibitors of Angiotensin Converting Enzyme. *Biochem. Biophys. Res. Commun.* 1985, *126*, 419–426.
(79) Godfrey, J. D.; Gordon, E. M.; Von Langen, D.; Engebrecht, J. Synthesis of Peptidyl Derived Amino Alcohols. *J. Org. Chem.* 1986, *51*, 3073.
(80) Rich, D. H.; Green, J.; Toth, M. V.; Marshall, G. R.; Kent, S. B. H. Hydroxyethylamine Analogs of the p17/p24 Substrate Cleavage Site are Tight-Binding Inhibitors of HIV Protease. *J. Med. Chem.* 1990, *33*, 1285–1288.
(81) Affymetrix Co. of Santa Clara, CA, a subsidiary of Affymax, uses the VLSIPS process for creation of oligonucleotide libraries on chips for the purposes of genetic typing and high-through-put DNA sequencing.

J. Am. Chem. Soc. 1992, 114, 10997-10998

$[Ag(CH_2Cl_2)_2]_2[Pd(OTeF_5)_4]$ (2.775 (2)-2.882 (2) Å).[10]

The most significant feature of the structure is the absence of Ag-O bonds, which are present in $[Ag(CH_2Cl_2)_2]_2[Pd(OTeF_5)_4]$,[10] $AgB(OTeF_5)_4$,[11b] and $[Ag(CO)][B(OTeF_5)_4]$.[11d] Instead, each $[Ag(CH_2Cl_2)_3]^+$ cation is only *extremely weakly coordinated* to the $Ti(OTeF_5)_6^{2-}$ anion by two Ag···F contacts of 3.029 (8) and 3.033 (6) Å. For comparison, the Ag-F distances in $AgSbF_6$[20] and AgF[21] are 2.62 and 2.467 (3) Å, respectively, and the sum of the van der Waals radii for silver and fluorine is 3.15 ± 0.08 Å.[22] The relative strength of anion binding to Ag^+ is also evident in the number of dichloromethane molecules coordinated to Ag^+—three in $[Ag(CH_2Cl_2)_3]_2[Ti(OTeF_5)_6]$ but only two in $[Ag(CH_2Cl_2)_2]_2[Pd(OTeF_5)_4]$.

In contrast with the $B(OTeF_5)_4^-$ anion,[11b] $Nb(OTeF_5)_6^-$ does not undergo rapid exchange with labeled $OTeF_5^-$ in the presence of electrophilic cations such as H^+ and Ag^+. For example, when $[TBA][Nb(^{16}OTeF_5)_6]$ and $H^{18}OTeF_5$ were mixed in dichloromethane at 22 °C, IR spectra showed that isotope scrambling was only 22% complete after 47 h. The presence of a larger cation had an even more dramatic effect: when $AgNb(^{16}OTeF_5)_6$ and $Ag^{18}OTeF_5$ were mixed in dichloromethane at 22 °C, *no* exchange was observed after 72 h. On the basis of the structure of $[Ag(CH_2Cl_2)_3]_2[Ti(OTeF_5)_6]$, we propose that electrophiles larger than H^+ cannot form bridge bonds to the *oxygen* atoms of $Nb(OTeF_5)_6^-$. Without such bridge bonds, abstraction of $OTeF_5^-$ by even the strongest electrophiles will not occur rapidly. Thus, steric hindrance causes a *kinetic* stabilization of $Nb(OTeF_5)_6^-$ (and presumably of other structurally related anions as well) in the presence of electrophilic cations.

Our new silver salts are freely soluble in weakly coordinating, low dielectric solvents such as chlorinated hydrocarbons and chlorofluorocarbons. For example, the solubility of $Ag_2Pd(OTeF_5)_4$ in dichloromethane at 22 °C ($\epsilon \approx 9.1$) is only 0.35 M,[10] while the solubility of $Ag_2Ti(OTeF_5)_6$ is *many* times higher (in fact, its solubility is sufficiently high that measuring it quantitatively has been problematic). An even more striking example of solubilizing ability is evident when comparing solubilities in CFC-113 at 22 °C ($\epsilon \approx 2.4$): $AgOTeF_5$, insoluble; $AgB(OTeF_5)_4$, 0.004 M; $AgNb(OTeF_5)_6$, 0.4 M.

The anions $Nb(OTeF_5)_6^-$ and $Ti(OTeF_5)_6^{2-}$ have the potential of being less coordinating, more stable in the presence of electrophilic cations, and more solubilizing than any previously reported anions. Detailed comparisons with anions such as $B(3,5\text{-}C_6H_3(CF_3)_2)_4^-$ and $CB_{11}H_{12}^-$ will be reported in a full article. The use of $Nb(OTeF_5)_6^-$, $Ti(OTeF_5)_6^{2-}$, and other very large, highly fluorinated anions for the preparation, isolation, and complete characterization of "coordinatively unsaturated" metal and metalloid cations remains an active endeavor in this laboratory.

Acknowledgment. We thank M. M. Miller for experimental assistance and the National Science Foundation for a grant in support of this research (CHE-9011610 to S.H.S.) and for grants to purchase the Siemens R3m diffractometer (CHE-8103011 to O.P.A.) and computing system (CHE-9107593 to O.P.A.). P. K.H. thanks the U.S. Department of Education for fellowship support under the Graduate Assistance in Areas of National Need Program (Grant No. P200A10210).

Supplementary Material Available: Tables S-I-VI, listing crystallographic data, atomic coordinates and isotropic thermal parameters, bond distances, bond angles, anisotropic thermal parameters, and hydrogen atom coordinates and thermal parameters (8 pages); Table S-VII, listing observed and calculated structure factors (10 pages). Ordering information is given on any current masthead page.

(20) Bode, H. *Z. Anorg. Allg. Chem.* **1951**, *267*, 62. An estimated standard deviation was not reported.

(21) Halleck, P. M.; Jarieson, J. C.; Pistorius, C. W. T. T. *J. Phys. Chem. Solids* **1972**, *33*, 769.

(22) (a) Bondi, A. *J. Phys. Chem.* **1964**, *68*, 441. (b) Pauling, L. *The Nature of the Chemical Bond*; Cornell University Press: Ithaca, NY, 1960; p 257.

A General and Expedient Method for the Solid-Phase Synthesis of 1,4-Benzodiazepine Derivatives

Barry A. Bunin and Jonathan A. Ellman*

Department of Chemistry, University of California Berkeley, California 94720

Received October 1, 1992

Very powerful methods have recently been developed for the combinatorial synthesis of large libraries of peptides which are then screened against a specific receptor or enzyme in order to determine the optimal peptide sequence for high affinity to that receptor or enzyme.[1] Unfortunately, peptides have limited utility as bioavailable therapeutic agents because they generally cannot be taken orally and have rapid physiological clearing times. The combinatorial synthesis and screening of bioavailable organic compounds would be a powerful extension of this approach. In this communication we report a general method for the expedient solid-phase synthesis of 1,4-benzodiazepine derivatives,[2] the critical first step in the combinatorial synthesis and screening of one of the most important classes of bioavailable therapeutic agents.[3] Because benzodiazepines are not polymers like the peptides and oligonucleotides that have previously been synthesized on solid support,[4] this report also demonstrates an important extension of solid-phase synthetic methods from the synthesis of biopolymers to the synthesis of nonpolymeric organic compounds.[5]

The 1,4-benzodiazepine derivatives are constructed on solid support from three separate components: 2-aminobenzophenones, amino acids, and alkylating agents (Scheme I). The 2-aminobenzophenone derivatives **1** are first attached to the polystyrene solid support through either a hydroxy or carboxylic acid functionality employing the acid-cleavable linker [4-(hydroxymethyl)phenoxy]acetic acid.[6] Synthesis of the benzodiazepine derivative on solid support then proceeds by removal of the FMOC protecting group from **2** by treatment with piperidine in DMF followed by coupling the resulting unprotected 2-aminobenzophenone to an α-*N*-FMOC-amino acid (Scheme I). Amide bond formation does not occur when the standard activation methods employed in solid-phase peptide synthesis are used (for example, carbodiimides and hydroxybenzotriazole or pentafluorophenyl active esters); however, treatment of the 2-aminobenzophenone with a methylene chloride solution of the α-*N*-FMOC-amino acid fluoride[7] in the presence of the acid scavenger 4-methyl-2,6-di-*tert*-butylpyridine results in complete coupling to provide amide **3**. The coupling conditions are suitable even for unreactive aminobenzophenone derivatives since complete coupling is observed for a derivative of **2** which contains both the *p*-chloro and the *m*-carboxy deactivating substituents (see **6i** and **6j** in Table I).

(1) (a) Fodor, S. P. A.; Read, J. L.; Pirrung, M. C.; Stryer, L.; Lu, A. T.; Solas, D. *Science* **1991**, *251*, 767-773. (b) Lam, K. S.; Salmon, S. E.; Hersh, E. M.; Hruby, V. J.; Kazmierski, W. M.; Knapp, R. J. *Nature* **1991**, *354*, 82-84. (c) Houghten, R. A.; Pinilla, C.; Blondelles, S. E.; Appel, J. R.; Dooley, C. T.; Cuervo, J. H. *Nature* **1991**, *354*, 84-86. (d) For a survey of these and other methods, see: Jung, G.; Beck-Sickinger, A. G. *Angew. Chem., Int. Ed. Engl.* **1992**, *31*, 367-383.

(2) Camps, F.; Cartells, J.; Pi *J. An Quim.* **1974**, *70(11)*, 848-9.

(3) Selected 1,4-benzodiazepine therapeutic agents or promising candidates: anxiolytics, hypnotics, and antiepileptics, Sternbach, L. H. *J. Med. Chem.* **1979**, *22*, 1-7; cholecystokinin antagonists, Evans, B. E.; et al. *J. Med. Chem.* **1988**, *31*, 2235-2246; opioid receptor antagonists, Romer, D.; et al. *Nature* **1982**, *298*, 759-760; HIV Tat antagonist, Hsu, M.-C. *Science* **1991**, *254*, 1799-1802; HIV reverse transcriptase inhibitor, Pauwels, R.; et al. *Nature* **1990**, *343*, 470-474; platelet activation factor antagonist, Miyazawa, S.; et al. *Chem. Pharm. Bull.* **1992**, *40* (2), 521-523 (thienodiazepine derivative).

(4) (a) Borany, G.; Merrifield, R. B. In *The Peptides*; Gross, E., Meienhofer, J., Eds.; Academic Press: New York, 1980; Vol. 2, pp 1-284. (b) Dorman, M. A.; Noble, S. A.; McBride, L. J.; Caruthers, M. H. *Tetrahedron* **1984**, *40*, 95-102.

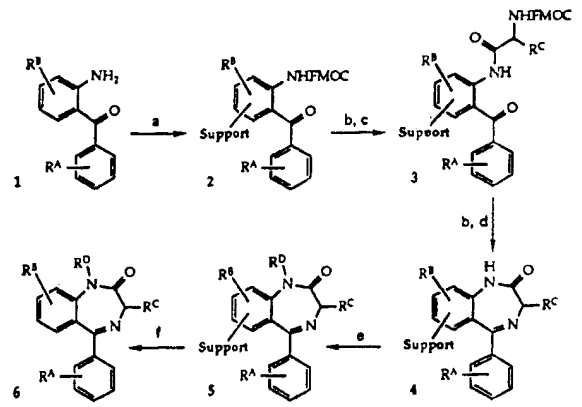
(5) Leznoff, C. C. *Acc. Chem. Res.* **1978**, *11*, 327-333.

(6) Sheppard, R. C.; Williams, B. J. *Int. J. Pept. Protein Res.* **1982**, *20*, 451-454. Full experimental details for coupling the 2-aminobenzophenone derivatives to the solid support are described in the supplementary material.

(7) Carpino, L. A.; Sadat-Aalaee, D.; Chao, H. G.; DeSelms, R. H. *J. Am. Chem. Soc.* **1990**, *112*, 9651-9652.

0002-7863/92/1514-10997$03.00/0

Scheme I[a]



[a] (a) See supplementary material; (b) 20% piperidine in DMF; (c) *N*-FMOC-amino acid fluoride, 4-methyl-2,6-di-*tert*-butylpyridine; (d) 5% acetic acid in DMF, 60 °C; (e) lithiated 5-(phenylmethyl)-2-oxazolidinone in THF, −78 °C, followed by alkylating agent and DMF; (f) $TFA/H_2O/Me_2S$ (95:5:10).

Table I. 1,4-Benzodiazepine Derivatives 6 (Scheme I)

entry	derivative R^A	R^B	R^C	R^D	yield (%)[a]
6a	4'-OH	5-Cl	CH_3	H	95
6b	4'-OH	5-Cl	CH_3	CH_3	100
6c	4'-OH	5-Cl	CH_3	CH_2CH_3	97
6d	4'-OH	5-Cl	CH_3	CH_2CHCH_2	90
6e	4'-OH	5-Cl	$CH(CH_3)_2$	CH_2CH_3	85
6f	4'-OH	5-Cl	CH_2CO_2H	CH_2CH_3	95
6g	4'-OH	5-Cl	$(CH_2)_4NH_2$	CH_2CH_3	95
6h	4'-OH	5-Cl	$CH_2Ph(4\text{-}OH)$	CH_2CH_3	98
6i		4-CO_2H,5-Cl	CH_2Ph	CH_3	100
6j		4-CO_2H,5-Cl	CH_3	CH_2Ph	93

[a] Yields are based on support-bound starting material 2.

The FMOC protecting group in 3 is then removed by treatment with piperidine in DMF. Exposure of the resulting free amine to 5% acetic acid in DMF provides the cyclic product 4. Complete cyclization is observed in the synthesis of 1,4-benzodiazepine derivatives with various steric and electronic properties (Table I), again demonstrating that general conditions have been identified for the solid-phase synthesis of diverse benzodiazepine derivatives.

Alkylation of the anilide of 4 then provides the fully derivatized 1,4-benzodiazepine 5 (Scheme I). Ideally, an excess of the base would be employed to achieve complete deprotonation and alkylation of the anilide, but employment of excess of commonly used bases such as LDA or NaH would result in deprotonation and subsequent alkylation of acidic functionality present elsewhere in the molecule. To maximize synthesis generality we therefore chose to employ lithiated 5-(phenylmethyl)-2-oxazolidinone[8] as the base since it is basic enough to completely deprotonate the anilide of 4, but not basic enough to deprotonate amide, carbamate, or ester functionalities. Upon deprotonation of 4, the appropriate alkylating agent is added followed by addition of anhydrous DMF to accelerate the alkylation reaction. By employing these conditions 1,4-benzodiazepine derivatives containing esters and carbamates have been alkylated in high yields on solid support with no overalkylation observed (compounds 6f and 6g in Table I where side chains were protected as a *tert*-butyl ester and a *tert*-butyl carbamate, respectively). Complete alkylation is observed for both activated alkylating agents such as methyl iodide and benzyl bromide and unactivated alkylating agents such as ethyl iodide.

The benzodiazepine product 5 is cleaved from the support with concomitant removal of acid-labile protecting groups by exposure to 85:5:10 trifluoroacetic acid/water/dimethyl sulfide. Employing this reaction sequence we have synthesized multiple structurally diverse benzodiazepine derivatives 6 in very high overall yields (Table I). In addition, racemization does not occur during the reaction sequence as determined by chiral HPLC analysis of the benzodiazepine derivatives 6a and 6c prepared from both (*R*)- and (*S*)-*N*-FMOC-alanine (Table I).[9] With the employment of this general and expedient solid-phase synthesis methodology, the construction and screening of a large combinatorial library of benzodiazepine derivatives are currently in progress and will be reported shortly. The solid-phase synthesis of other classes of therapeutically important organic compounds is also under investigation and will be reported in due course.

Supplementary Material Available: Listings of experimental procedures for attaching the aminobenzophenone derivatives to the solid support and for the solid-phase synthesis of the benzodiazepine derivatives, including analytical data for all of the 1,4-benzodiazepine derivatives and intermediates (8 pages). Ordering information is given on any current masthead page.

(8) The pK_a of 5-(phenylmethyl)-2-oxazolidinone is 20.5 in DMSO as determined by Bordwell. Evans, D. A.; et al. *J. Am. Chem. Soc.* **1990**, *112*, 4011-4030. Lithiated 5-(phenylmethyl)-2-oxazolidinone is employed rather than unsubstituted 2-oxazolidinone due to its greater solubility in tetrahydrofuran.

(9) Pirkle, W. H.; Tsipouras, A. *J. Chromatogr.* **1984**, *291*, 291-298.

Direct Evidence for an Oxocarbenium Ion Intermediate in the Asymmetric Cleavage of Chiral Acetals

Tarek Sammakia* and Randall S. Smith

Department of Chemistry and Biochemistry
University of Colorado
Boulder, Colorado 80309-0215

Received September 4, 1992

The Lewis acid mediated cleavage of chiral acetals has been the subject of numerous investigations over the past several years and is a useful tool for the asymmetric synthesis of carbon–carbon bonds.[1] With allylsilanes[2] and allylstannanes[3] as nucleophiles, this reaction can provide chiral ethers with diastereoselectivities ranging from 5:1 to >500:1. The origin of this selectivity has been studied, mostly by studying the behavior of model acetals,[4] but these studies are inconclusive because the behavior of the model compounds is known to vary with minor variations in the structure of the acetals.[5] We wish to report the results of a more direct approach to studying the reactivity of chiral acetals which utilizes the stereospecifically deuterated acetal 1[6] (Table I) to determine

(1) Johnson, W. S.; Harbert, C. A.; Stipanovic, R. D. *J. Am. Chem. Soc.* **1968**, *90*, 5279. Johnson, W. S.; Harbert, C. A.; Ratcliffe, B. E.; Stipanovic, R. D. *J. Am. Chem. Soc.* **1976**, *98*, 6188. McNamara, J. M.; Kishi, Y. *J. Am. Chem. Soc.* **1982**, *104*, 7371. Sekizaki, H.; Jung, M.; McNamara, J. M.; Kishi, Y. *J. Am. Chem. Soc.* **1982**, *104*, 7372. For a recent review of the chemistry of chiral acetals, see: Alexakis, A.; Mangeney, P. *Tetrahedron: Asymmetry* **1990**, *1*, 477.

(2) Johnson, W. S.; Crackett, P. H.; Elliot, J. D.; Jagodzinski, J. J.; Lindel, S. D.; Natarajan, S. *Tetrahedron Lett.* **1984**, *25*, 3951.

(3) Denmark, S. E.; Almstead, N. G. *J. Org. Chem.* **1991**, *56*, 6485.

(4) For previous mechanistic studies of this type of acetal opening, see: (a) Bartlett, P. A.; Johnson, W. S.; Elliott, J. D. *J. Am. Chem. Soc.* **1983**, *105*, 2088. (b) Choi, V. M. F.; Elliot, J. D.; Johnson, W. S. *Tetrahedron Lett.* **1984**, *25*, 591. (c) Mori, A.; Fujiwara, J.; Maruoka, K.; Yamamoto, H. *J. Organomet. Chem.* **1985**, *285*, 83. (d) Maruoka, K.; Yamamoto, H. *Angew. Chem., Int. Ed. Engl.* **1985**, *24*, 668. (e) Yamamoto, Y.; Nishi, S.; Yamada, J. *J. Am. Chem. Soc.* **1986**, *108*, 7116. (f) Silverman, R.; Edington, C.; Elliot, J. D.; Johnson, W. S. *J. Org. Chem.* **1987**, *52*, 180. (g) Murata, S.; Suzuki, M.; Noyori, R. *Tetrahedron* **1988**, *44*, 4259. (h) Schreiber, S. L.; Wang, Z. *Tetrahedron Lett.* **1988**, *29*, 4085. (i) Denmark, S. E.; Willson, T. M.; Almstead, N. G. *J. Am. Chem. Soc.* **1989**, *111*, 9258. (j) Mori, I.; Ishihara, K.; Flippin, L. A.; Nozaki, K.; Yamamoto, H.; Bartlett, P. A.; Heathcock, C. H. *J. Org. Chem.* **1990**, *55*, 6107. (k) Denmark, S. E.; Almstead, N. G. *J. Am. Chem. Soc.* **1991**, *113*, 8089. (l) Denmark, S. E.; Almstead, N. G. *J. Org. Chem.* **1991**, *56*, 6458. (m) Sammakia, T.; Smith, R. *J. Org. Chem.* **1992**, *57*, 2997.

(5) See refs 4j-m.

0002-7863/92/1514-10998$03.00/0

EXHIBIT G

Proc. Natl. Acad. Sci. USA
Vol. 90, pp. 6909–6913, August 1993
Chemistry

"Diversomers": An approach to nonpeptide, nonoligomeric chemical diversity

(multiple synthesis/solid-phase synthesis/hydantoin/benzodiazepine)

SHEILA HOBBS DEWITT*, JOHN S. KIELY†, CHARLES J. STANKOVIC, MEL C. SCHROEDER, DONNA M. REYNOLDS CODY, AND MICHAEL R. PAVIA‡

BioOrganic Chemistry Section, Department of Chemistry, Parke-Davis Pharmaceutical Research Division, Warner-Lambert Co., 2800 Plymouth Road, Ann Arbor, MI 48106

Communicated by Pedro Cuatrecasas, April 8, 1993

ABSTRACT **Solid-phase chemistry, organic synthesis, and an apparatus for multiple, simultaneous synthesis have been combined to generate libraries of organic compounds ("diversomers"). Arrays of compounds were synthesized over two to three steps incorporating chemically diverse building blocks on a polystyrene-based solid support in a multiple, simultaneous manner. The generality of this approach is illustrated by the syntheses of dipeptides, hydantoins, and benzodiazepines.**

Recent advances in robotics, miniaturization, and automation have resulted in the development of rapid, high throughput biological screening assays, which can quickly exhaust available sources of chemical diversity. Driven by these advances in biological testing, several methods of generating chemical diversity, primarily peptide- or nucleotide-based oligomer libraries, have been developed using multiple, simultaneous chemical synthesis (1–4) or molecular biology techniques (5–7).

While undeniably useful in chemical lead discovery, these oligomer libraries are limited. The chemical leads discovered from these libraries still require extensive modifications to produce suitable drug candidates due to the general unsuitability of peptides or oligonucleotides as stable, orally active drugs. Furthermore, the available chemical fragments, or building blocks, are generally limited, even allowing for the use of unnatural enantiomers or artificial amino acids and modified nucleotides. Finally, these oligomeric libraries contain a repetitive backbone linkage, either amides or phosphates, which is contrary to the concept of diversity.

Recently, Simon *et al.* (8) reported the development of oligomeric N-substituted glycines as motifs for the generation of compound libraries. Although these libraries offer several advantages over peptide or oligonucleotide libraries—for example, different structural variations, reduced susceptibility to hydrolysis, and the incorporation of achiral building blocks—they still suffer from a repetitive amide motif. One method for the synthesis of nonpeptide, nonoligomeric compounds was recently reported by Bunin and Ellman (9) for the solid-phase synthesis of 1,4-benzodiazepine derivatives. Their method, however, is limited by requisite introduction of auxiliary functionality (e.g., hydroxy or carboxylic acid) in the target molecule to facilitate attachment to the solid support. Furthermore, no apparatus or method for multiple, simultaneous synthesis is described.

True chemical diversity is only achievable by removing the restrictions on the bond-forming reactions and building blocks used. The resulting compound libraries would have the same advantages and uses as oligomer-based libraries with the added benefit that the target structures are limited only by the chemists' creativity.

We have developed an apparatus and method for the multiple, simultaneous synthesis of organic compounds to create unique collections of compounds, which we call "diversomers." This apparatus and method greatly increases the flexibility and diversity of structures that can be produced by multiple, simultaneous synthesis technology. The target compounds are simultaneously, but separately, synthesized on a solid support in an array format (Fig. 1) to generate an ensemble of structurally related compounds. Using this diversomer method, we have simultaneously synthesized and characterized 40 discrete hydantoins and 40 discrete benzodiazepines. The benzodiazepines were then screened in a competitive binding assay to ascertain biological activity.

MATERIALS AND METHODS

Apparatus. An apparatus suitable for multiple, simultaneous synthesis on a solid support is represented in Fig. 2. The key feature of the apparatus is the use of a gas dispersion tube, which we refer to as a "pin," to physically contain the solid support (resin) during the reactions. This concept of "resin *in* a pin" is contrasted to current solid-phase multiple, simultaneous synthesis technology, which contains the resin in a reservoir well (10, 11) ("resin *in* a well") or on a polypropylene rod (12) ("resin *on* a pin"). Our design overcomes several disadvantages including limited amounts of final products (<<1 mg), the inability to produce soluble products, difficulties in manipulating solid supports, and, in particular, the general unsuitability for normal organic synthesis techniques. The fritted glass filters serve to contain the solid support, allow efficient mixing between reactants in the reservoir wells and the resin in the frit, and facilitate separation of the resin-bound intermediates from excess reagents, solvents, and by-products. The holder block serves to secure the pins and provides a means for simultaneously manipulating the array as a single unit. The reservoir block is constructed with an array of reaction wells that accommodate the fritted glass filters, while concurrently retaining a quantity of reactant necessary to perform the required reactions. Separate reaction wells allow individual reactions to be executed and monitored, while maintaining the integrity of filtrates, intermediates and final products corresponding to each location in the array. The manifold encloses the upper portion of the pins, allowing control over the reaction atmo-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: Fmoc, 9-fluorenylmethoxycarbonyl; ISTD, internal standard; TFA, trifluoroacetic acid; DM. , *N,N*-dimethylformamide; Boc, *t*-butyoxycarbonyl; SAR, structure–activity relationship; meq, milliequivalent(s).

*To whom reprint requests should be addressed.
†Present address: ISIS Pharmaceuticals, Carlsbad, CA 92008.
‡Present address: Sphinx Pharmaceuticals, Cambridge, MA 02139.

Building Block #1	Building Block #2 A	B	C	A	B	C	A	B	C
1	1AX	1BX	1CX	1AY	1BY	1CY	1AZ	1BZ	1CZ
2	2AX	2BX	2CX	2AY	2BY	2CY	2AZ	2BZ	2CZ
3	3AX	3BX	3CX	3AY	3BY	3CY	3AZ	3BZ	3CZ
	X			Y			Z		

Building Block #3

FIG. 1. Two sets of building blocks (1–3 and A–C) undergo a single bond forming reaction to give nine elements (molecules). A third set of building blocks (X–Z) and a second reaction will provide 27 elements.

sphere. Extended pins can serve as condensers when a chilled gas is circulated through the manifold, thus providing a means to maintain reflux. The apparatus is secured with gaskets and clips to create a gas-tight unit. Reagents can be added directly into the pins by injection through a gasket-sealed plate at the top of the manifold. Due to the heterogeneous nature of these reactions, agitation is critical to success and can be achieved by rotational platform shaking, magnetic stirring, or preferably by sonication.

Automation. A Tecan model 5032 robotic sample processor (Tecan U.S., Research Triangle Park, NC) was used to perform all of the liquid sample handling and TLC spotting.

General Synthesis. 9-Fluorenylmethoxycarbonyl (Fmoc)-protected amino acid *p*-alkoxybenzyl alcohol resins [Wang resins (13), 200–400 mesh, 1% crosslinked with divinylbenzene, 0.37–0.60 milliequivalent (meq)/g], hydroxymethyl resin (200–400 mesh, 1% crosslinked with divinylbenzene, 1.04 meq/g), and *t*-butoxycarbonyl (Boc)-protected amino acid Merrifield resins (200–400 mesh, 1% crosslinked with divinylbenzene, 0.57–0.89 meq/g) were obtained from Bachem. Other chemicals and solvents were obtained from Aldrich and EM Science and were used without further purification. All synthetic compounds and intermediates gave satisfactory ^{1}H NMR and MS. Gas chromatography was performed on a Varian model 3400 instrument.

The standard-array synthesis cycle begins with weighing an appropriate amount of solid support into each pin and fitting the pins into the holder block equipped with gaskets. The manifold and the reaction wells are attached, and a positive flow of nitrogen is maintained throughout the total reaction sequence. The resins are swelled in the first reaction

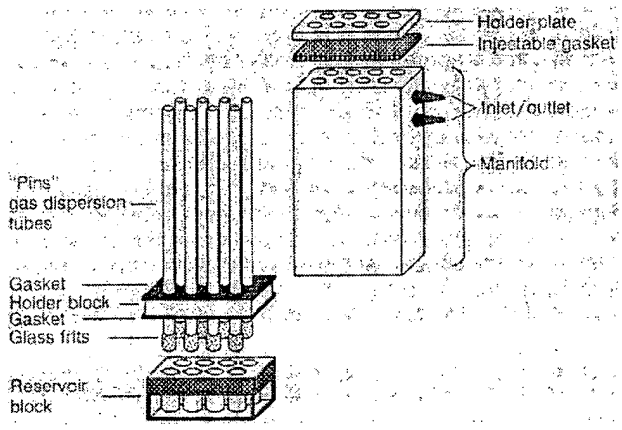


FIG. 2. The apparatus consists of an array of gas dispersion tubes (pins), a reservoir block with multiple reaction wells, a holder block, a manifold, and gaskets.

solvent and are drained by gravity in preparation for the first reaction cycle. Reagents are dispensed, as appropriate, into clean reaction wells or directly into the pins through the gasket-sealed plate at the top of the manifold. The apparatus is heated, cooled, and agitated as necessary to complete or equilibrate the reactions. Reaction times are determined from preliminary validation studies and/or monitoring of filtrates during the course of the reaction [GC or HPLC in conjunction with internal standard (ISTD) calibration methods]. At the end of each reaction the pins are subjected to a series of wash cycles to remove residual solvents, reagents, and by-products. A standard wash protocol includes sequentially submerging the pins in 2 × 3–5 ml each of *N,N*-dimethylformamide (DMF), methanol, water, water/dioxane (1:1), dioxane, and CH_2Cl_2, followed by agitation by sonication for 10–15 min, and finally draining by gravity and nitrogen purge. The efficiency of the wash cycles is monitored by TLC or GC/ISTD of the wash filtrates.

Synthesis of Hydantoins. Forty discrete hydantoins were synthesized by deprotecting, then treating each of eight amino acid resins with each of five isocyanates. This three-step synthesis creates an array with eight variations in step two and five variations in step three to generate a 40-unit (8 × 5) two-dimensional array (Fig. 3, data not shown).

Approximately 95–105 mg each of seven Fmoc-protected amino acid Wang (13) resins (phenylalanine, glycine, isoleucine, leucine, alanine, valine, and tryptophan) and 95–105 mg of Boc-protected diphenylglycine (prepared by condensation with hydroxymethyl resin) was measured into 40 pins. The resins were swollen with 3 ml of DMF.

To deprotect the Fmoc-amino acids, the appropriate pins were submerged in reaction wells containing 3 ml each of 25% (vol/vol) piperidine/DMF with an ISTD (anthracene). To deprotect the Boc-amino acids, the appropriate pins were submerged in reaction wells containing 3 ml each of 50% (vol/vol) trifluoroacetic acid (TFA)/DMF. The apparatus was agitated for 6 hr in a sonic bath, and the reaction progress was monitored by analyzing aliquots of the filtrates for the Fmoc–piperidine adduct and dibenzofulvene by GC/ISTD calibration methods.

Resin-bound ureas were synthesized by submerging the appropriate pins in reaction wells containing 3 ml each of the appropriate isocyanate in DMF (5–20 eq), with an ISTD (anthracene). The apparatus was agitated in a sonic bath for 6 hr, and the reaction was monitored by derivatization of a sample of the filtrate with an appropriate amine or alcohol, followed by analysis by GC/ISTD calibration methods.

To cyclize and cleave the final hydantoins from the solid support, the pins were submerged in reaction wells containing 3 ml each of 6 M aqueous HCl. The apparatus was heated in an oil bath at 85–100°C for 2 hr, while maintaining a positive chilled nitrogen flow through the manifold.

To isolate and purify the products, the pins were submerged in reaction wells, each containing 3 ml each of methanol. The apparatus was agitated in a sonic bath for 10–15 min to extract the hydantoins from the resins. The methanol extraction protocol was repeated three times, until the filtrates were free of any organic components, as determined by TLC. The HCl and methanol filtrates were concentrated on a Speed-Vac instrument to afford 39 of the 40 desired hydantoins as the HCl salts. A total of 0.3–11.5 mg corresponding to 4–81% yield of the desired hydantoins was isolated and analyzed by TLC, MS, and ^{1}H NMR.

Synthesis of Benzodiazepines. Forty discrete benzodiazepines were synthesized by treating each of five amino acid resins with each of eight 2-amino benzophenone imines. This two-step synthesis created an array with five variations in step one and eight variations in step two to generate a 40-unit (8 × 5) two-dimensional array (Fig. 4 and Table 1).

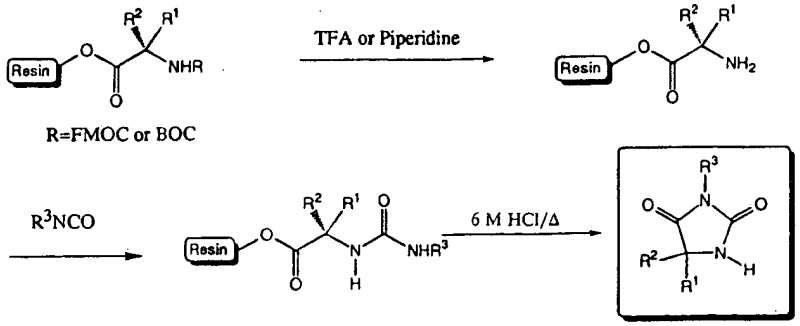


FIG. 3. General route for synthesis of hydantoins.

Five Boc-protected amino acid Merrifield resins (alanine, glycine, phenylalanine, tryptophan, and valine) were deprotected in bulk (1–5 g), using TFA/CH_2Cl_2, 1:1, at 25°C overnight. The resins were washed with dioxane and CH_2Cl_2, then dried under vacuum, and used directly as the TFA salts.

Approximately 99–107 mg each of five-amino acid Merrifield resins prepared above was measured into 40 pins. The resins were swollen with 3 ml each of CH_2Cl_2.

To form the resin-bound imines, the pins were submerged in reaction wells containing 3 ml each of the appropriate 2-aminobenzophenone imines (3–6 eq) in dichloroethane and heated at 60°C (oil-bath temperature) for 24 hr. The pin array was then drained and washed by repeatedly dispensing 4-ml portions of CH_2Cl_2 through the aperture at the top of each pin 12 times, until the washes were no longer colored.

To cyclize and cleave the benzodiazepines from the solid support, the pins were submerged in reaction wells containing 3 ml each of 100% TFA. The apparatus was submerged in an oil bath (60°C) and heated for 20 hr.

To isolate and purify the final products, the pins were drained and then extracted by repeatedly dispensing 2-ml portions of CH_2Cl_2 through the aperture at the top of each pin three times. The combined washes and reaction filtrates were evaporated under a stream of nitrogen. A simple two-phase extraction procedure was implemented using the Tecan 5032 processor. The residues from evaporation were dissolved in 3 ml of CH_2Cl_2 and mixed with 3 ml of saturated aqueous $NaHCO_3$. The organic phase was withdrawn, and the aqueous layer was extracted twice more with 1.5 ml of CH_2Cl_2. The combined organic extracts were dried with $MgSO_4$, filtered, and concentrated as before to yield the expected benzodiazepines. The 40 products were characterized by TLC, ^1H NMR, and MS. A total of 2–14 mg corresponding to 9–63% yield of each benzodiazepine was isolated with estimated purities (by ^1H NMR) typically >90%.

Assay for Inhibition of Fluoronitrazepam. The assay performed by NovaScreen (Scios-Nova Pharmaceutical, Baltimore) employed bovine cortical membranes with [^3H]fluoronitrazepam as the radioligand. The percent inhibition of radioligand binding was determined at three concentrations (10^{-9}, 10^{-7}, and 10^{-5} M). IC_{50} data were calculated for each compound from the average of two determinations at each concentration.

RESULTS AND DISCUSSION

Central to the demonstration of the diversomer library is the development of a "general" method for the multiple, simultaneous synthesis of organic molecules. This method should satisfy the following criteria. The compounds should be simultaneously, but separately, synthesized in an array format (Fig. 1) in a soluble form and in sufficient quantity (>1 mg) and purity to allow multiple *in vitro* biological testing. The apparatus should be compatible with all the normal techniques of organic synthesis. To the extent possible, the method should use automation for speed, accuracy, and precision. Finally, the intermediate products should be readily separable from by-products and excess reagents. To satisfy this last requirement we chose to implement solid-phase synthesis techniques using functionalized, cross-linked polystyrene resins. The overall feasibility of this approach was suggested by literature reports for the synthesis of nonpeptide molecules on resins (14–16).

In the diversomer method, building blocks are sequentially coupled to the growing molecule on the resin until the penultimate, resin-bound product at each location in the array is complete. Cleavage from the resin yields a final product that can be readily separated from the spent resin. Several options for cleavage are illustrated in Fig. 5. The preferred strategy (cleavage 1) constructs a resin-bound penultimate product possessing distal functionality which, when activated or unmasked, will attack the resin-linking bond and eject the cyclized product into solution. Because compounds that do not "react" remain attached to the resin, this option should provide purer final products.

Validation Studies. Initially, a diversomer synthesis requires verification of the proposed resin-based synthesis including several representative examples before an array synthesis is performed. Fourier transform-infrared, enhanced with deconvolution techniques (17–19), and ^{13}C gel-phase NMR (20–22) of the resin-bound intermediates provide an excellent means to directly monitor the progress of reactions on a polystyrene support. These techniques enable qualitative assessment of the success of reactions on a solid support but are not amenable for use within the apparatus. (Their use within the apparatus would require the removal of a prohibitive amount of resin from the pins.) However, analysis of the reaction filtrates by GC or HPLC in conjunc-

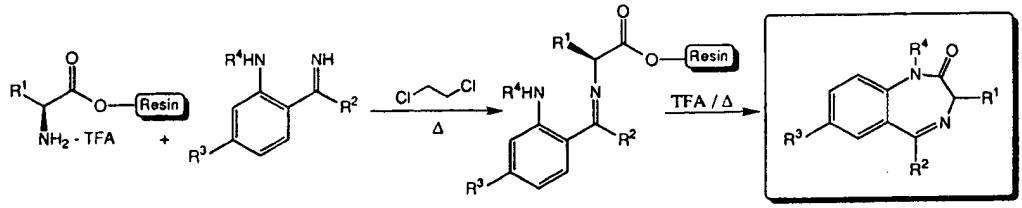
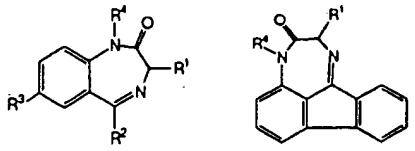


FIG. 4. General route for synthesis of benzodiazepines.

Table 1. Benzodiazepines generated in array



Structure A Structure B

No.	R^1*	R^2*	R^3	R^4	Yield†, mg	Yield†, %	IC_{50}‡
1	Me	Ph	H	H	6.1	40	1,700 nM
2	Me	Ph	Cl	H	9.6	56	200 nM
3	Me	4-MeOPh	H	H	5.8	34	69,000 nM
4	Me	Ph	NO_2	H	4.9	28	91 nM
5	Me	B§	B§	H	9.6	63	29,000 nM
6	Me	Ph	Cl	Me	3.2	18	160 nM
7	Me	Chx	H	H	6.4	41	31,000 nM
8	Me	2-Thn	H	H	7.4	47	5,500 nM
9	H	Ph	H	H	9.4	44	1,100 nM
10	H	Ph	Cl	H	13.7	55	19 nM
11	H	4-MeOPh	H	H	5.5	23	33,000 nM
12	H	Ph	NO_2	H	8.0	31	16 nM
13	H	B§	B§	H	3.4	16	44,000 nM
14	H	Ph	Cl	Me	5.2	20	21 nM
15	H	Chx	H	H	7.0	32	6,100 nM
16	H	2-Thn	H	H	8.8	41	940 nM
17	Bzl	Ph	H	H	8.6	52	19,000 nM
18	Bzl	Ph	Cl	H	8.8	46	1,800 nM
19	Bzl	4-MeOPh	H	H	7.3	41	>100 μM
20	Bzl	Ph	NO_2	H	4.9	26	2,400 nM
21	Bzl	B§	B§	H	8.6	52	>100 μM
22	Bzl	Ph	Cl	Me	2.5	13	5,000 nM
23	Bzl	Chx	H	H	6.5	39	>100 μM
24	Bzl	2-Thn	H	H	8.4	48	47,000 nM
25	3-MeInd	Ph	H	H	9.5	43	69,000 nM
26	3-MeInd	Ph	Cl	H	8.0	33	16,000 nM
27	3-MeInd	4-MeOPh	H	H	7.4	31	>100 μM
28	3-MeInd	Ph	NO_2	H	5.8	23	12,000 nM
29	3-MeInd	B§	B§	H	5.2	23	>100 μM
30	3-MeInd	Ph	Cl	Me	2.5	10	14,000 nM
31	3-MeInd	Chx	H	H	7.8	34	>100 μM
32	3-MeInd	2-Thn	H	H	9.2	40	71,000 nM
33	iPr	Ph	H	H	7.1	31	>100 μM
34	iPr	Ph	Cl	H	7.0	28	>100 μM
35	iPr	4-MeOPh	H	H	7.1	29	>100 μM
36	iPr	Ph	NO_2	H	2.2	9	>100 μM
37	iPr	B§	B§	H	6.4	29	>100 μM
38	iPr	Ph	Cl	Me	3.0	11	82,000 nM
39	iPr	Chx	H	H	6.0	27	>100 μM
40	iPr	2-Thn	H	H	8.4	37	>100 μM

*Bzl, benzyl; 3-MeInd, 3-methylindole; 4-MeOPh, 4-methoxyphenyl; Chx, cyclohexyl; 2-Thn, 2-thienyl; iPr, isopropyl.
†Yields are based on indicated loading of commercially available functionalized resins (0.50–0.89 meq/g).
‡Approximate IC_{50} values are based on three-point fit. Values were also obtained for the commercially available diazepam (1.46 nM), nordiazepam (0.2 nM), and nitrazepam (0.67 nM), corresponding to sample numbers 14, 10 and 12, respectively.
§See structure B.

tion with an ISTD provides an indirect method to monitor the course of the individual reactions within the array. This technique provides a means to quantitatively monitor the uptake or release of reagents from the resin.

On the basis of preliminary experiments in the validation stage and consideration of reaction kinetics (e.g., steric or electronic contributions), the reaction conditions are selected to drive the slowest or poorest synthesis in the array to completion.

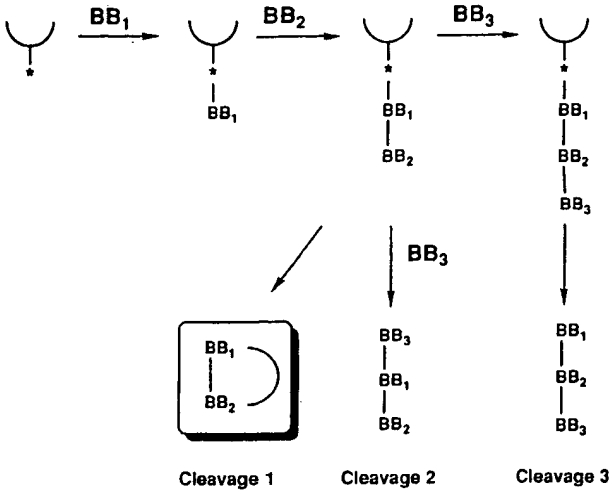


FIG. 5. BB_n represents building blocks that are bifunctional to allow sequential attachment. The curved structure represents the solid support, where the star is the functionality capable of covalently attaching the growing molecule to the solid support. The three possible cleavage modes are illustrated.

The well-established nature of resin-based peptide synthesis (23–25) provided an ideal means to validate the ability of our apparatus to successfully execute multiple, simultaneous reactions. Thus, several arrays of dipeptides were synthesized using standard Fmoc chemistry (26) and a prototype 8-pin apparatus (data not shown). Each product was analyzed by HPLC, MS, and ^1H NMR. The yields (30–85%) and purities were within our defined specifications and, thus, verified the utility of the apparatus to produce the expected and desired products as designed.

Demonstration of Diversomer Method. Although the synthesis of dipeptides serves to demonstrate the utility of the apparatus, it is insufficient to validate the general approach. To validate the diversomer method, we needed to simultaneously synthesize an array of compounds over several synthetic steps, which would produce a representative portion of the known structure–activity relationship (SAR) around a pharmaceutically relevant structure. The syntheses of a series of hydantoins and benzodiazepines were chosen for this demonstration because of the large SAR already developed around the commercial products Dilantin and Valium, respectively.

Synthesis of Hydantoins. The synthesis of an array of 40 hydantoins, including Dilantin, was achieved using the 40-pin apparatus. Five groups of eight Fmoc- or Boc-protected amino acid resins were deprotected and then separately treated with eight groups of five isocyanates, followed by treatment with aqueous 6 M HCl to generate 40 discrete hydantoins (Fig. 3, data not shown).

Some of the key features that demonstrate the strengths, flexibility, and scope of the method and apparatus are as follows. (*i*) The 40 Boc- or Fmoc-protected resin-bound amino acids were simultaneously deprotected. (*ii*) The equilibration or completion of both the Fmoc deprotection and isocyanate reactions was monitored by GC/ISTD calibration methods. (*iii*) The final cleavage and extractions were efficiently performed even in nonswelling solvents. (*iv*) Residual reactants or by-products were removed by wash cycles that included submersion and sonication of the pins in a series of solvents. (*v*) Wash cycles were monitored by robotic spotting of the filtrates on a TLC plate and observing the results under UV light to insure the removal of residual reagents and byproducts.

Synthesis and Testing of Benzodiazepines. The investigation of an array synthesis of benzodiazepines as potential targets was prompted by a brief communication by Camps *et al.* (27), who reported a one-step synthesis of several benzodiazepines starting with a resin-bound amino acid. The synthesis of an array of 40 benzodiazepines related to Valium was achieved by using the 40-pin apparatus. Eight groups of five-amino acid resins were trans-imidated (28, 29) with five groups of eight 2-aminobenzophenone imines (30, 31) to form resin-bound imines, followed by treatment with TFA (32) to generate 40 discrete benzodiazepines (Fig. 4 and Table 1).

To verify that the compounds produced could be used directly in a biological assay, the crude benzodiazepines were assayed, *without* further purification, for inhibition of fluoronitrazepam (Table 1). The compounds expected to exhibit the greatest activity based on the known SAR (33) were the most potent compounds in the array [e.g., R^1 = methyl (**1–8**) or hydrogen (**9–16**) and R^3 = chloro (**2, 6, 10, 14**) or nitro (**4, 12**)]. The data also provide sufficient semiquantitative information to permit reasonable conclusions concerning the SAR. For example, substituents larger than methyl at R^1 are not tolerated, electron-withdrawing substituents at R^3 are favored, and R^2 appears to be limited to aromatic rings.

Conclusions and Uses. The value of the diversomer approach is represented in the simplicity of the concept, the timeliness for drug discovery efforts, and the innovative combination of methods and apparatus to demonstrate the multiple, simultaneous synthesis of nonpeptide, nonoligomeric compounds on a solid support. The feasibility of preparing arrays of compounds has been demonstrated by the multiple, simultaneous synthesis of 40 discrete hydantoins and 40 discrete benzodiazepines. This result was accomplished with an apparatus capable of performing 40 simultaneous, but separate, chemical reactions on a solid support. The chemistry compatible with the diversomer method encompasses nearly all organic reactions, and the apparatus is sufficiently general and complete to allow for most organic synthesis techniques. The generality and compatibility of this apparatus represents immense improvements over current equipment for multiple, simultaneous synthesis on a solid support (10–12). For example, temperature control, agitation, inert atmosphere, injection of sensitive reagents, and reaction monitoring are an integral part of the apparatus. The key feature of this method is the "resin *in* a pin" apparatus design, which provides a means to simultaneously segregate and manipulate an array of resin-bound intermediates. Furthermore, the simplicity of our apparatus design facilitates the ease of use and construction from commercially available components in a variety of dimensions and multiplicities.

The diversomer method greatly increases the flexibility and diversity of structures that can be produced by multiple, simultaneous synthesis technology. Although the number of compounds produced in a single array (40 separate compounds) is significantly smaller than that which can be prepared as mixtures by some of the current methods for generating peptide libraries (10^5–10^7 peptides), this is adequately offset by the increased quantities and purities of products and the enhanced chemical and structural diversity that can be achieved using the diversomer approach. For example, a typical molecular mass range for orally available drugs is 500–600 g/mol; within this range there are approximately 2.3 million possible tetrapeptides (using 39 D- and L-amino acids), whereas there are an unlimited number of nonpeptide structures. Because each molecule is produced separately and can be fully characterized by standard analytical techniques, no time-consuming deconvolution techniques are needed to determine the active constituent in large mixtures of compounds. Additionally, the quantities produced are sufficient to allow screening in multiple *in vitro* assays.

The diversomer library provides an array of compounds that are directly suitable for biological testing, thereby dramatically enhancing and accelerating SAR development by the ability to screen multiple, related compounds at one time. Well-designed arrays of compounds will provide information necessary for lead compound discovery and SAR evaluation. With continued use, diversomers will provide unlimited sources of chemical diversity.

We are grateful to W. Moos for his leadership in initiation of this research and guidance during its conceptualization, J. Topliss for his long-term advocacy of this type of chemical diversity, and J. Bristol for his continued support. We express our deep gratitude to R. Root-Bernstein for teaching us a new way of looking at research problems and to C. Agree, A. Galen, D. Holsworth, D. Mack, H.-Y. Mei, D. Moreland, and K. Sanders for helpful discussions. We are indebted to R. Harms and D. Van Hofe for their expert craftsmanship in constructing the apparatus.

1. Dower, W. J. & Fodor, S. P. A. (1991) *Annu. Rep. Med. Chem.* **26**, 271–280.
2. Fodor, S. P. A., Read, J. L., Pirrung, M. C., Stryer, L., Lu, A. T. & Solas, D. (1991) *Science* **251**, 767–773.
3. Jung, G. & Beck-Sickinger, A. G. (1992) *Angew. Chem. Int. Ed. Engl.* **31**, 367–383.
4. Zuckermann, R. N., Kerr, J. M., Siani, M. A., Banville, S. C. & Santi, D. V. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 4505–4509.
5. Scott, J. K. & Smith, G. P. (1990) *Science* **249**, 386–390.
6. Devlin, J. J., Panganiban, L. C. & Devlin, P. E. (1990) *Science* **249**, 404–406.
7. Cwirla, S. E., Peters, E. A., Barrett, R. W. & Dower, W. J. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 6378–6382.
8. Simon, R. J., Kania, R. S., Zuckermann, R. N., Huebner, V. D., Jewell, D. A., Banville, S., Ng, S., Wang, L., Rosenberg, S., Marlowe, C. K., Spellmeyer, D. C., Tan, R., Frankel, A. D., Santi, D. V., Cohen, F. E. & Bartlett, P. A. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 9367–9371.
9. Bunin, B. A. & Ellman, J. A. (1992) *J. Am. Chem. Soc.* **114**, 10997–10998.
10. Gausepohl, H., Kraft, M., Boulin, C. & Frank, R. W. (1990) in *Peptides: Chemistry, Structure, and Biology*, Proceedings of the 11th American Peptide Symposium, eds. Rivier, J. & Marshall, G. (ESCOM, Leiden, The Netherlands), pp. 1003–1004.
11. Schnorrenberg, G. & Gerhardt, H. (1989) *Tetrahedron* **45**, 6031–6040.
12. Geyson, H. M., Meloen, R. H. & Barteling, S. J. (1984) *Proc. Natl. Acad. Sci. USA* **82**, 5131–5135.
13. Wang, S.-S. (1973) *J. Am. Chem. Soc.* **95**, 1328–1333.
14. Leznoff, C. C. (1978) *Acc. Chem. Res.* **11**, 327–333.
15. Crowley, J. I. & Rapoport, H. (1976) *Acc. Chem. Res.* **9**, 135–144.
16. Leznoff, C. C. (1974) *Chem. Soc. Rev.* **3**, 65–85.
17. Larsen, B. D., Holm, A., Christensen, D. H., Werner, F. & Nielsen, O. F. (1992) in *Innovation and Perspectives in Solid Phase Synthesis*, ed. Epton, R. (Intercept, Andover, U.K.), pp. 363–366.
18. Kauppinen, J. K., Moffatt, D. J., Mantsch, H. H. & Cameron, D. G. (1981) *Anal. Chem.* **53**, 1454–1457.
19. Byler, D. M. & Susi, H. (1986) *Biopolymer* **25**, 469–487.
20. Mapelli, C. & Swerdloff, M. D. (1990) in *Peptides*, eds. Giralt, E. & Andreu, D. (ESCOM, Leiden, The Netherlands), pp. 316–319.
21. Epton, R., Goddard, P. & Ivin, K. J. (1980) *Polymer* **21**, 1367–1371.
22. Giralt, E., Rizo, J. & Pedroso, E. (1984) *Tetrahedron* **40**, 4141–4152.
23. Merrifield, R. B. (1963) *J. Am. Chem. Soc.* **85**, 2149–2154.
24. Bayer, E. (1991) *Angew. Chem. Int. Ed. Engl.* **32**, 113–129.
25. Stewart, J. M. & Young, J. D., eds. (1984) *Solid Phase Peptide Synthesis* (Pierce Chemical, Rockford, IL).
26. Atherton, E., Gait, E. J., Sheppard, R. C. & Williams, B. J. (1979) *Bioorg. Chem.* **8**, 351–370.
27. Camps, F., Cartells, J. & Pi, J. (1974) *An. Quim.* **70**, 848–849.
28. Schmitt, J., Suquet, M., Boitard, J., Comoy, P., Meingan, J. P., Gallet, G., Clim, T., Perrin, C. & LeMeur, J. (1967) *Ind. Chim. Belge.* **32**, 184–187.
29. O'Donnell, M. J. & Polt, R. L. (1982) *J. Org. Chem.* **47**, 2663–2666.
30. Meguro, K. & Kuwada, Y. (1972) U.S. Patent 3,587,941.
31. Sternbach, L. H. & Richen, G. S. (1967) U.S. Patent 3,297,755.
32. Sugasawa, T., Adachi, M., Sasakura, K., Matsushita, A., Eigyo, M., Shiomi, T., Shintaku, H., Takahara, Y. & Murata, S. (1985) *J. Med. Chem.* **28**, 699–707.
33. Sternbach, L. H., Randall, L. O., Banziger, R. & Lehr, H. (1968) in *Drugs Affecting the Central Nervous System*, ed. Burger, A. (Dekker, New York), pp. 237–264.

EXHIBIT H

Pharmaceutical Research, Vol. 8, No. 2, 1991

Review

Mass Receptor Screening for New Drugs

Ronald M. Burch[1,2] and Donald J. Kyle[1]

Mass receptor screening is capable of identifying drug candidates in large compound libraries. Our laboratory has developed a mass screening technology by standardizing assay protocols that can be transferred from receptor to receptor. The entire operation, from disbursement of compounds to data analysis, is computerized to handle vast numbers of experimental results. The success of this method depends upon strict definitions of compound activity, with rapid elimination of compounds that do not fulfill all criteria. Finally, we approach automation with caution. While certain items, such as automatic harvesters, are essential for high-throughput screening, much time can be spent optimizing gadgets instead of gathering data.

KEY WORDS: receptor; data analysis; ligand binding; mass screening; HyperCard; structure data base.

INTRODUCTION

Mass ligand binding screening or "receptor screening" to detect receptor agonists or antagonists is a relatively recent technique. Many new peptide neurotransmitters, peptide and protein hormones, cytokines, and growth factors are tempting therapeutic targets. In the case of smaller peptides, little secondary structure is apparent in solution (1), and deletion of certain residues may result in inability of the analogue to bind to its receptor. The deleted residue may not play a role in binding; instead, it may normally serve as part of the "messenger portion" of the molecule, directing it to the proper level in the membrane to enable the "address portion" actually to bind to and activate the receptor (2). In the case of larger protein ligands, secondary and tertiary structure exist, but few structures are available, and at this time it is not understood what parts of the molecules interact with receptors. For example, interleukin I exists as two different molecules, interleukin 1α and interleukin 1β, each with a molecular mass of 17,000 daltons. Both bind to the receptor with a similar affinity, yet they share only 26% sequence homology (3). The recent explosion in molecular cloning of receptors has revealed largely unsuspected diversity in what were thought to be well-described receptor classes. For example, at least five distinct subtypes of muscarinic acetylcholine receptors have been identified (4), while pharmacological probes had firmly detected only two (5). The existence of additional receptor subtypes, often with distinct distributions, makes it certain that far more specific therapeutic agents may be possible. Examination of old libraries of compounds may yield agents quite specific for the new receptors.

Mass screening strategies may be *directed*, that is, specific compounds may be chosen, or the strategy employed may be *random* among available compounds in chemical or natural product libraries. Nonpeptide analogues of peptide agonists have been derived from natural products, for example, opiates and cholecystokinin antagonists (6). Our laboratory can screen 10,000 compounds per month in an assay with a team of only three technicians. However, reliance on ligand binding assays without validation of leads in biological assays, can lead down synthetic blind alleys (7). While ligand-binding methodology results in few false negatives, certain assays generate many false-positive results. This review addresses ligand binding assays for high-throughput screening.

THEORY OF LIGAND BINDING ASSAYS

The theoretical basis of ligand–receptor binding assays has been previously described (8,9). The dissociation constant, K_d, with units of moles per liter, is used here as a measure of affinity. Many physiological ligands and therapeutic agents interact with their receptors in such assays with K_d values of a few nanomolar to as low as a few picomolar, while few therapeutic agents exist with K_d values in the micromolar range.

Receptor affinities are determined by incubating multiple tubes with identical amounts of tissue preparation and radioligand but with increasing amounts of the test compound. IC_{50} values of tracer displacement are calculated as shown in Fig. 1. To account for tracer receptor binding, the Cheng–Prusoff equation (10) can be used to calculate the K_d from the IC_{50} of the test compound for the receptor, the value usually being written "K_i" to denote that it was obtained by the competition method just described.

The Ideal Assay

The success of the binding assay depends upon two

[1] Nova Pharmaceutical Corporation, 6200 Freeport Centre, Baltimore, Maryland 21224.
[2] To whom correspondence should be addressed.

0724-8741/91/0200-0141$06.50/0

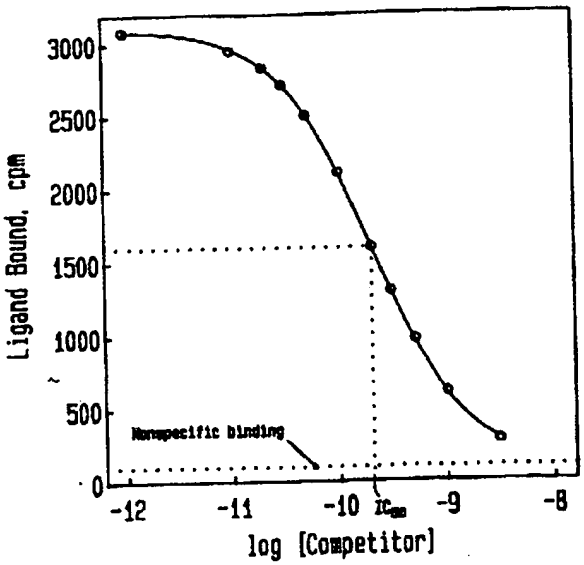


Fig. 1. "Competition curve" for a test compound to compete with a radiolabeled ligand for binding at a receptor. Total binding is 3100 cpm. Nonspecific binding, 100 cpm, was obtained in the presence of 1 μM unlabeled ligand (not shown). Thus, specific binding to the receptor is 3100 − 100 = 3000 cpm. One-half of total binding is 1500 cpm (plus 100 nonspecific), or 1600 cpm. The dotted line shows interpolation of the $IC_{50} = 10^{-9.7}$ M, or 200 pM. From the IC_{50}, the K_i value can be calculated (see Ref. 10).

components, the tissue and the ligand. The aims in choosing the tissue and ligand are the attainment of a high signal-to-noise ratio and the specificity of the assay.

Tissue. An ideal tissue has a high density of receptors in order to achieve a high signal without the need for large amounts of tissue (which often leads to high noise because of difficulty in separating unbound ligand). However, many of the receptors for recently discovered biologically active peptides are found in low abundance. For example, interleukin 1 receptors are found at densities as low as 200 per cell in lymphocytes (11). In the case of interleukin 1, use of other tissue sources in screening is possible. Fibroblasts can express 3000–5000 interleukin 1 receptors per cell (11), while cells "commonly" express 15,000 to 50,000 receptors for many hormones. In addition, certain transformed lymphocytes express tens of thousands of receptors per cell (12).

Suppose one wishes to identify antagonists of the lymphocyte interleukin 1 receptor. The low number of receptors expressed by normal lymphocyte would make their use too costly. One may choose fibroblasts or a transformed line, but it must then be shown that the binding sites on these cells are identical to the target receptor on normal lymphocytes. For the interleukin 1 receptor these problems have been solved; molecular cloning has demonstrated that the receptors on normal lymphocytes, transformed lymphocytes, and fibroblasts all have identical structures (12,13).

One must also be aware that a *binding site* is not necessarily a *receptor*. Many cells express B2 bradykinin receptors. Binding sites were identified in several tissues, including guinea pig ileum, murine and human fibroblasts, and neuroblastoma cells (14), which were affected in predictable ways by known bradykinin analogs. However, one "receptor" identified on neuroblastoma cells was later shown to actually represent a metabolic enzyme, angiotensin converting enzyme (15).

Another potential source of misinformation is the presence of more than one subtype of receptor in a preparation. This is of particular concern when complex tissues such as brain are used, and in receptor systems for which agonists must be used as ligands (see below under *Ligand*). The presence of multiple receptor subtypes in a tissue may result in confusing results, or, if one subtype is present as a small percentage of the total, it may be missed entirely. Several possible methods can be used to reduce the problem of receptor heterogeneity. Radioligands specific for only a single receptor subtype may be used. However, specific ligands are rarely available.

A novel approach is the use of cloned receptors, stably expressed in cells which do not usually express any binding site for the ligand that is used (16). Such systems provide unambiguous assays, using *human* receptors without the need for access to human tissues.

Ligand. An antagonist is the ideal ligand. In many systems agonists may not recognize all receptors (17). Also, many of the binding assays for cytokines and growth factors utilize intact cells. Unless care is taken to use very low assay temperature, many agonists will be internalized as complexes with their receptors (18), making unreliable any "binding parameters" obtained. Finally, a radiolabeled endogenous agonist will rarely discriminate among receptor subtypes, since nature intended for all to recognize the ligand, and virtually all natural ligands have similar affinities for receptors (100 pM–5 nM). Unfortunately, when screening against newly described receptors, antagonists are rarely available; identification of the first one is often the goal of the screening exercise!

Mass Screening Protocols

Mass ligand binding screening requires enormous planning and coordination. Following are examples of how screening efforts are coordinated in our laboratory. The first is appropriate for small-scale projects; the other is suitable for large projects.

Discovery of a New Lead for a Single Receptor

Manual Approach. A single technician is required to harvest the tissue, prepare it, obtain samples of test compounds from the compound disbursement facility, set up and terminate the assay, prepare the filtered samples for radioactivity counting, calculate results, and add them to a data base. Using this protocol, technicians usually perform assays 2 days per week. Assays are set up in 48- or 96-tube racks or in 96-well plates. Either twelve 48-tube racks, or six 96-tube racks are set up per day, resulting in 576 tubes. Each rack contains duplicate tubes for total bound counts and nonspecifically bound counts. Each day a K_i is determined for a reference compound. Thus, about 520 tubes are available for test compounds. Since each compound is assayed using a single tube, 520 different compounds are assayed per day.

Tissues, buffers, and ligands are prepared in the morning. Incubation requires 1–3 hr, followed by termination of

the assay using a cell harvester. All 576 tubes can be filtered and washed within 15 min. The filters are then punched into counting vials.

The next day, the K_i of the reference compound and percentage inhibition of binding by the test compounds are calculated, then entered into a data base. Hard copies of all the raw count data and calculations may be affixed in notebooks. A single binding technician can screen about 1000 compounds per week. Each technician in the compound room can weigh and solubilize about 250 compounds per day. Thus, one disbursement technician is required per binding technician. A library of 10,000 compounds requires about 10 weeks for initial activity determination, using a total of 20 technician-weeks.

High-Throughput Approach. Screening of a large library through a single receptor assay is better performed by assay teams, composed of one technician who performs the binding assay and two technicians who disburse solubilized compounds. The binding technician performs a 576-tube assay every day, so that in a week's time, about 2500 test compounds can be screened. Such high throughput depends critically on computerization.

All compounds are identified with bar code labels. Molecular weights and other pertinent information are entered into the computer (Fig. 2). Balances are interfaced to the computer. Thus, to disburse a compound, a technician passes a bar code reader over the vial, opens the vial, and places an aliquot of compound onto the balance pan. The computer reads the mass and prints a label identifying the compound number and the volume of solvent (usually dimethyl sulfoxide) to add to the sample to reach the desired stock concentration. The technician places the sample into a vial, affixes the label, and adds solvent. The next sample may then be processed. At the end of the disbursement process (520 compounds plus a reference compound) the computer prints a table containing compound identification numbers, disbursement numbers, amount weighted, molecular weight, and volume of solvent added to reach a stock concentration. This information is passed on to the binding technician with the samples.

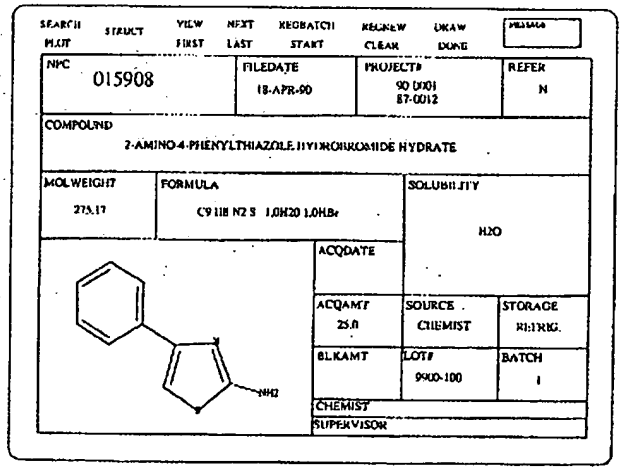


Fig. 2. Nova Pharmaceutical Corporation compound submission data record form.

The computer acquires data from the counters in real time. The next day, the binding technician is required to spend only a few minutes to call up the raw data to assure that total and nonspecific binding are within standard limits; then a calculation program is chosen along with the identifying number of the file of compounds in the assay. The computer performs all calculations, then, upon approval of the technician, enters data automatically into a data base. With such a system the binding technician is relieved of the tedium of data calculation and entry, and the system is not corrupted by data entry errors. Entered into the notebook are the assay protocol (these are standard and are entered only as a number), the number of the file containing the raw data, and the list of compounds that were assayed.

Using this technology, a 10,000-compound library requires only 4 weeks for determination of activity, and only 12 technician-weeks is required.

Simultaneous Screening at Multiple Receptors

Most pharmaceutical companies possess large libraries of compounds. These libraries are valuable resources, for they may contain the prototypes for new therapeutic classes of drugs. Every year, new receptors are discovered and described in the scientific literature. Thus, a chemical library may be rescreened year after year, in new assays.

In a multiple screening paradigm, integrated teams of disbursement and assay technicians are not required, since a single disbursement will serve for many different assays. When such a project is under way, the separate groups of disbursements are stored at 4°C in a central location.

For each assay, 4 weeks is required, as described above. In Fig. 3 another consideration is illustrated, the "low-signal" assay. To this point, the projected time-lines and labor estimates have assumed assays of "high signal." A high signal assay has little nonspecific binding, for example, our bradykinin assay, with binding of 98%. Total binding might be 3000 cpm, while nonspecific binding is about 100 cpm. If an active compound is one that inhibits binding by 50%, then few false-positives or -negatives will occur based on counting errors. However, certain assays, for example those for eicosanoids, have only about 50% specific binding. Assuming total binding of 2000 cpm, then nonspecific bind-

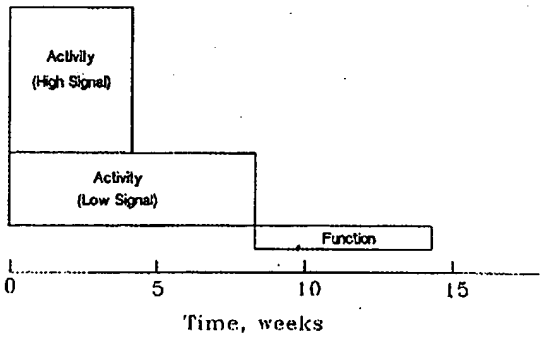


Fig. 3. Time-line for screening a library of 10,000 compounds in 30 different receptor assays, 20 with a high signal-to-noise ratio and 10 with a low signal-to-noise ratio (the relative heights of the boxes). The time assumes a single binding technician for each assay, all performed simultaneously.

ing will be 1000 cpm. An active compound will reduce binding to 1500 cpm. Clearly, with a window of 500 cpm, or one-third to one-fourth of the total binding, there will inevitably be a significant number of false-positives or -negatives. Thus, in low-signal assays, all compounds are screened twice, extending the time required for determination of initial activity to 8 weeks. All compounds that are inactive in one trial and active in the other are rescreened together a third time. In Fig. 3, we assumed 30 separate receptors, 20 high signal and 10 low signal. Disbursement of compounds requires 4 technician-weeks. The 20 high-signal assays require 80 technician-weeks; the 10 low-signal assays, 80 technician-weeks.

Secondary Binding Screening

After initial activity testing, K_i values are determined for all active compounds (Fig. 3). Two compounds are assayed per 48-tube rack, a single technician performing 24 determinations per day. Ideally, K_i values are determined using fresh disbursements of compounds. Potency testing requires very little time compared to activity testing (Fig. 3).

In addition to determining potency in the assay in which a compound is active, all compounds are tested for specificity, often in 30–40 different receptor binding assays. Generally, this is done using a single concentration of the compound, with secondary potency testing being performed in assays in which the compound was active. Of course, when initial screening of a library is done in a battery of assays (Fig. 3), specificity testing in binding is built into the initial screening effort.

FUNCTIONAL SCREENING

After a compound has been found active and specific, its agonist or antagonist properties are determined, since binding assays cannot distinguish between the two. Functional assays can range from second-messenger assays to properties at isolated tissues. Second-messenger assays are chosen appropriate to the receptor type being studied, for example, cAMP accumulation, calcium mobilization, prostaglandin synthesis, and inositol phosphate formation. Functional assays should be performed in systems that are maximally complex without sacrificing too much in speed. We often choose isolated smooth muscle preparations. They are often sensitive to poorly specific compounds, exhibiting increased irritability or depressed responsiveness, making detection of nonspecific compounds less difficult.

In functional screening several concentrations of the compound are tested to determine whether it elicits the effect expected of an agonist. Next, its ability to inhibit the effect of a known agonist is determined. Any compound that exhibits a "negative" activity *must* be tested to determine whether it is acting at a specific receptor, or exerting some nonspecific or toxic effect. A negative effect refers to inhibition of some process. For example, bradykinin stimulates inositol phosphate formation in fibroblasts (20). Thus, a functional assay for detection of a bradykinin antagonist might consist of determining whether a test compound can block bradykinin-induced inositol phosphate formation. Inhibition of the process may take place, not only at the level of the bradykinin receptor, but also at the level of an enzyme in the pathway to inositol phosphate or at the level of cell viability. When using negative assays, any active compound must be tested carefully for specificity. For the assay just described, specificity testing might consist of eliciting inositol phosphate formation with thrombin or bombesin, which act at their own, distinct receptors. The test compound, if it acts at the bradykinin receptor, should have no effect.

If possible, assays should be designed such that positive effects are elicited by active test compounds. For example, tumor necrosis factor causes cytotoxicity (21). A positive assay for a tumor necrosis factor receptor antagonist might consist of determining whether a test compound can block the cytotoxic effect of tumor necrosis factor.

DEFINING SUCCESS

We define "active" from 50% inhibition of binding at a concentration of test compound of 1 m*M* to 70% inhibition of binding at 1 μM. The definition may depend upon the "hit rate" of the assay. In excitatory amino acid binding assays, hit rates may be as high as 10–20% at test compound concentrations of 10 μM; in interleukin 1 binding assays the hit rate is 1 in 5000 or less. If a high concentration of test compound is required to detect activity, it is less likely that a specific interaction is taking place between the compound and a receptor (7).

Useful lead compounds have had K_i values in binding assays no higher than a few micromolar: our own initial bradykinin antagonist lead, NPC 361, K_i of 400 n*M* (14); the initial Merck cholecystokinin antagonist lead, asperlicin, K_i of 0.6 μM (6); and the Dupont angiotensin II antagonist lead, K_i of 40 μM (22). Most often our "hit criterion" is 50% inhibition at a 10 μM concentration of test compound, this assuring a K_i of no more than 10 μM.

Of importance equal to potency is *specificity*. In the past we have synthetically pursued "leads" with K_i's of 1 μM that were nonspecific in functional assays. In no case has any useful compound been developed (7). An impotent compound or a potent, nonspecific compound is not a viable lead.

PROCESS TECHNOLOGY

We have found that most mechanical automation techniques do *not* increase throughput. Ligand binding assays are very straightforward. The use of reservoir-equipped pipettes, such as the Eppendorf Combitip, is enormously more rapid than using robot systems. Adding either radioligand or tissue to 576 tubes requires approximately 5 min using a Combitip. A robot requires at least as much time, plus programming, plus extensive maintenance.

High-throughput assays are impossible without an automatic sample filtration device. Brandel harvesters simultaneously filter and wash 48 samples using a single filter mat, in about 1 min. Heads are available to fit several formats, such as 48- or 96-tube racks and 96-well plates. Each filter port has a surface area of 250 mm^2; filtration area is a very important consideration when using solid scintillants (19).

Filling and capping scintillation vials are time-consuming. Now, however, filter mats containing solid scintillant are available (19), which require no filling or capping of tubes. Counting efficiency using the Brandel apparatus

with solid scintillant is comparable to liquid scintillants. Use of filtration apparati with smaller surface areas is associated with dramatically reduced counting efficiency due to "tissue stacking" (19).

COMPUTERIZATION: THE KEY TO HIGH THROUGHPUT

Inventory Considerations

Computerization of every possible step in the binding laboratory, from disbursement of compounds to collecting and analyzing data, is the cornerstone of maximizing throughput. Our central VAX facility is accessed via terminal emulation, through VTERM (Coefficient Systems, New York) on "IBM-compatable" computers and VersaTerm-Pro (Synergy Software, Reading, Pennsylvania) on Macintosh computers. Each compound is assigned an identification number that corresponds to the "parent structure" (i.e., the molecular structure as it would exist in the salt-free form). Its source, molecular weight, salts, information on solubility, and literature references are stored on a data record form (Fig. 2) that is translated into a database using the Customization Module of the MACCS II program. The interface serves as a window into two separate data-base systems housed on the VAX computer ORACLE (Oracle Corporation, California) for inventory and biological data and MACCS II (Molecular Design Ltd., San Leandro, California) for chemical structures. Several modes of error checking are used, including duplicate checking and molecular weights. Molecular structure "drawing rules" are followed while drawing a chemical structure in MACCS II, to provide guidelines for the visual orientation of a chemical structure going into MACCS II, so that upon retrieval in report format, there will be uniformity between them.

Managing the Biological Data

There are two approaches available for translating biological results into the central data-base system. Since mass screening assays are composed of multiple racks of identical arrangement, it is preferable to transmit raw data directly from the radioactivity counters to the VAX. Using a set of C programs, the output from any given counter is routed through a VAX RS-232 port, then written to a data file of text format resident on the VAX. From the data file, percentage inhibitions, IC_{50}'s, K_i's, and Hill coefficients are determined. To enhance the review process, a listing of only those compounds considered as "hits" on the basis of a predetermined percentage inhibition is also generated. Furthermore, the results from those samples added to the rack as quality control are listed independently, since their biological profile is known in advance and a quick analysis of these calculated biological values can provide insight to the integrity of the experiment. After the review process, the technician is able to order the VAX to dump the results directly into the centralized ORACLE tables for subsequent public access.

Biological data can also be entered onto the VAX manually. Data from functional assays in most cases tends to be "customized," thereby making the more automated systems nongeneric and overly complex.

The MacIntosh is used as the platform for developing a circumvention of the "traditional VT 100 type" data entry procedure (SQL*MENU or SQL*PLUS). This required both the ORACLE program and Apple's HyperCard. Using a custom interface developed within the HyperCard environment, a technician is able to enter data manually, in free format, into a scrolling HyperCard field. Sorting data, numbering experiments, extensive error checking, printing of hard copies, and writing out text files containing the data are functions handled through the Macintosh interface. In addition, the cut, copy, and past features on the Macintosh are used to simplify data entry. The SQL*LOADER facility is used to fill ORACLE tables resident on the Macintosh with the text files written by scientists during the previous week. Subsequently, these tables are transferred to the VAX over the ethernet using the "copy table" command in ORACLE. Individually tailored hard-copy reports may be generated that display structures, and whatever data are appropriate. In effect there is no direct interaction between the scientists and the centralized data-base facility.

For those instances where the scientists must access the VAX ORACLE tables, another HyperCard card is used to structure complex SQL queries (Fig. 4). Thus, staff members are able to insert, update, or simply view the centralized ORACLE data easily with little training in SQL queries. Implementation of this technology relies on Ethernet boards attached to the Macintosh computers and SQL*NET networking protocol running on both PCs and the VAX computer. A comprehensive overview of the flow of electronic information is presented in Fig. 5.

BEYOND LEAD IDENTIFICATION

Computerized Structural Search Paradigms for Pharmacophore Identification

The ability to search binding data in as many as 70–100 assays and functional data in dozens of assays is a powerful tool when coupled to the ability to search chemical libraries by structure. Molecular structures are electronically stored in a variety of formats, usually dictated by some combination of atom type and a connectivity scheme. Data-base pro-

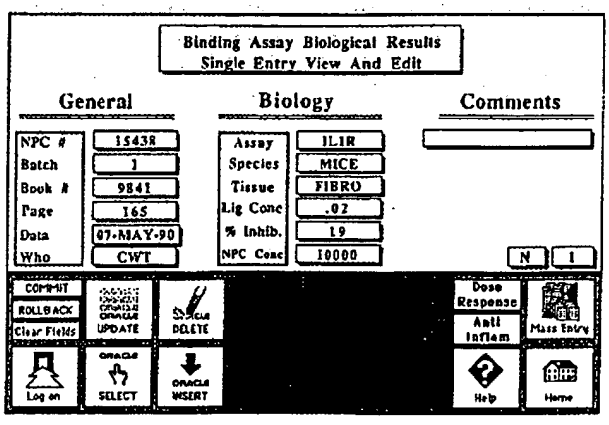


Fig. 4. HyperCard based form used for structuring complex SQL queries. The form accesses the centralized VAX:ORACLE tables.

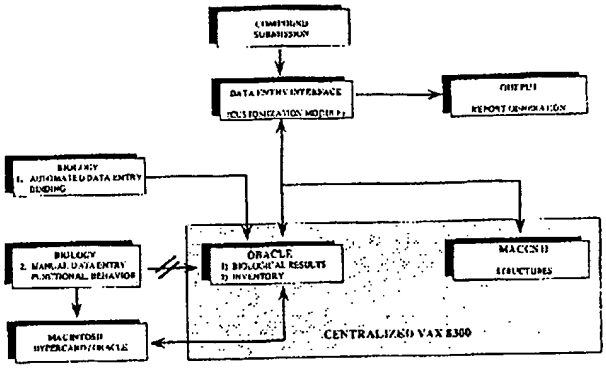


Fig. 3. Overview of the flow of information electronically at Nova Pharmaceutical Corporation.

grams that handle structures generally provide utilities for initiating any one of a variety of searches. Examples include searching by molecular formula, chemical name, or substructure, the latter being of most interest to the practicing medicinal chemist. The substructure itself may be either jointed or disjointed, completely defined atom by atom and bond by bond, or variable. Consideration of these substructures is possible either alone or, in some cases, in association with other physical characteristics including pK, spectroscopic data, or partition coefficients if those data are available.

The primary goal of the medicinal chemist is to establish a relationship between the three-dimensional structure of a series of molecules and their measured biological activities (SAR, structure–activity relationship). Upon formulation of an SAR, the next step is the design of chemical entities, consistent with the hypothesis, that are expected to show an enhancement in the biological property if the SAR is valid. Formulation of an SAR and the subsequent preparation and biological testing of these molecules become a repetitive cycle that ideally can guide the chemist toward a structural entity with desirable therapeutic properties.

The structural data base in combination with the binding and functional assay information can be used to search for features that include or exclude certain structures from being active. Several reports have appeared describing the application of similarity and dissimilarity measures to the storage and retrieval of structural information. In one example (23), classification of local anesthetics according to similarity and dissimilarity coefficients between pairs of structure diagrams and application of cluster analysis to the results was similar to biologic classification.

Perhaps of more significance is the application of similarity and dissimilarity measures as an enhancement to random receptor binding screening programs. The similarity measure of a compound is represented as a vector of chemical descriptors in chemical descriptor space. Association of the biological activity of a molecule with that chemical descriptor space has utility in defining a similarity–activity space. Hence, lead compound discovery might be reduced to locating a compound in a new region of this space. Lead compound optimization, by analogy, might be viewed as locating a structure within the same region of space as its parent but that represents a move to a more active geographical site. Selection of compounds for subsequent receptor screening on the basis of dissimilarity has been proposed as an effective means of locating new lead structures by obtaining the widest sampling of similarity–activity space, outside of that region defined by the initial receptor hits, with a minimal number of compounds.

Extensions of the two-dimensional structural data bases based on connectivity are the three-dimensional structural data bases capable of storing multiple conformations of any given chemical structure together with their corresponding physical properties, either measured or calculated [MACCS 3D, MENTHOR (24): CHEMSTAT (Chemical Design Ltd., Oxford, England)]. These data bases are ideal end points for housing conformations derived from molecular modeling studies, X-ray crystallography, or NMR experiments. Furthermore, the information is handled in three dimensions, completely consistent with the logic of the chemists destined to make use of the information. Some of the issues surrounding the practical implementation of a three-dimensional data base include how many conformations for each molecular structure will be stored and the method of gradient convergence, whether those conformations are determined by molecular mechanics, semiempirical methods, or *ab initio* methods, and whether the structures are local energy minima or dynamic. Methods are available for rapidly converting two-dimensional chemical structures into three-dimensional molecular coordinates [CONCORD (Evans and Sutherland Computer Corp., Salt Lake City, Utah); CORBA (Oxford Molecular, Oxford, England)]. Each of these methods is constructed upon expert systems, eliminating the need for any numerical evaluation of either a wave function or a classical potential energy expression. There are also growing numbers of commercial data bases that contain the three-dimensional coordinates of selected groups of molecular structures.

Although there is much uncertainty as to the most effective implementation of a three-dimensional data-base system, the most promising applications will likely be related to three-dimensional searches based on some interesting pharmacophore pattern. Unlike the substructure searches run on two-dimensional molecular structures where "matches" are effectively predefined on the basis of the connectivity of the substructure, the three-dimensional search has the capability of matching the relative spatial orientations of functional groups or atoms irrespective of the connectivity between them. In searching a subset of the Cambridge Crystallographic Database for those molecules which could fit inside the combined volume of several known nicotinic agonists and which had interatomic distances compatible with a given pharmacophore geometry (25), several novel designs for nicotinic agonists were derived. In another example, ALADDIN was used to test alternative superposition rules for mapping of the D2 dopamine receptor, then design compounds to fit the known binding site. Indeed, three compounds were discovered in the search that had activity at the D2 receptor (26).

Computer technology is providing powerful tools with direct application to drug discovery programs. Theoretical properties for molecules can be calculated and saved in association with a chemical structure. All can be stored electronically as part of a centralized database system containing

diverse biological data. Although in its infancy now, this technology should minimize duplication of synthetic efforts on the part of medicinal chemists and should also provide a wealth of information in support of, and advancement of, their ongoing structure–activity relationship hypotheses.

ACKNOWLEDGMENT

We wish to thank Cheryl Sowards for preparing the manuscript.

REFERENCES

1. D. J. Kyle, R. P. Hicks, P. R. Blake, and V. J. Klimkowski. Conformational properties of bradykinin and bradykinin antagonists. In R. M. Burch (ed.), *Bradykinin Receptor Antagonists: Basic and Clinical Research*, Marcel Dekker, New York, 1990, pp. 131–146.
2. R. Schwyzer. A new principle in QSAR: Membrane requirements. *J. Receptor Res.*, in press (1990).
3. U. Gubler, A. O. Chua, A. S. Stern, C. P. Hellmann, M. P. Vitek, T. M. Dechiara, W. R. Benjamin, K. J. Collier, M. Dukovich, P. C. Familletti, C. Fiedler-Nagy, J. Jenson, K. Kaffka, P. L. Kilian, D. Stremlo, B. H. Wittreich, D. Woehle, S. B. Mizel, and P. T. Lomedico. Recombinant human interleukin 1 alpha: Purification and biological characterization. *J. Immunol.* 136:2492–2497 (1986).
4. N. J. M. Birdsall and E. C. Hulme. Muscarinic receptor subclasses. *Trends Pharmacol. Sci.* 4:459–463 (1983).
5. T. I. Bonner. The molecular basis of muscarinic receptor diversity. *Trends Neurosci.* 12:148–151 (1989).
6. R. S. L. Chang, V. J. Lotti, R. L. Monaghan, J. Birnbaum, E. O. Stapley, M. A. Goetz, G. Albers-Schonberg, A. A. Patchett, J. M. Liesch, O. D. Hensens, and J. P. Springer. A potent nonpeptide cholecystokinin antagonist selective for peripheral tissues isolated from *Aspergillus alliaceus*. *Science* 230:177–180 (1985).
7. R. M. Burch. Mass ligand binding screening for receptor antagonists: Prototype new drugs and blind alleys. *J. Receptor Res.* in press (1990).
8. L. E. Limbird. *Cell Surface Receptors: A Short Course on Theory and Methods*, Martinus Nühoff, Boston, 1986, p. 196.
9. G. A. Weiland and P. B. Molinoff. Quantitative analysis of drug-receptor interactions. I. Determination of kinetic and equilibrium properties. *Life Sci.* 29:313–330 (1981).
10. Y. C. Cheng and W. H. Prusoff. Relationship between the inhibition constant (K_i) and the concentration of inhibitor which causes 50% inhibition (I_{50}) of an enzymatic reaction. *Biochem. Pharmacol.* 22:3099–3108 (1973).
11. S. K. Dower, S. M. Call, S. Gillis, and D. L. Urdal. Similarity between the interleukin 1 receptors on a murine T-lymphoma cell line and on a murine fibroblast cell line. *Proc. Natl. Acad. Sci. USA* 83:1060–1064 (1986).
12. J. E. Sims, C. J. March, D. Cosman, M. B. Widmer, H. R. MacDonald, C. J. McMahan, C. E. Grubin, J. M. Wignall, J. L. Jackson, S. M. Call, D. Friend, A. R. Alpert, S. Gillis, D. L. Urdal, and S. K. Dower. cDNA expression cloning of the IL-1 receptor, a member of the immunoglobin superfamily. *Science* 241:585–589 (1988).
13. J. E. Sims, R. B. Acres, C. E. Grubin, C. J. McMahan, J. M. Wignall, C. J. March, and S. K. Dower. Cloning the interleukin 1 receptor from human T cells. *Proc. Natl. Acad. Sci. USA* 86:8946–8950.
14. S. G. Farmer and R. M. Burch. Pharmacology of bradykinin receptors. In R. M. Burch (ed.), *Bradykinin Receptor Antagonists: Basic and Clinical Research*, Marcel Dekker, New York, 1990, pp. 1–31.
15. J. G. de Vries, E. Phillips, C. R. Snell, P. H., Snell, and M. Webb. Construction of a physiologically active photoaffinity probe based on the structure of bradykinin; Labelling of angiotensin converting enzyme but not candidate bradykinin receptor NG108-15 cells. *J. Neurochem.* 52:1508–1516 (1989).
16. N. J. Buckley, T. I. Bonner, C. M. Buckley, and M. R. Brann. Antagonist binding properties of five cloned muscarinic receptors expressed in CHO—K1 cells. *Mol. Pharmacol.* 35:469–476 (1989).
17. A. DeLean, J. M. Stadel, and R. J. Lefkowitz. A ternary complex model explains the agonist-specific binding properties of the adenylate cyclase-coupled β-adrenergic receptor. *J. Biol. Chem.* 255:7108–7117 (1980).
18. R. J. Lefkowitz, M. G. Carson, and G. L. Stiles. Mechanisms of membrane-receptor regulation. *N. Engl. J. Med.* 310:1570–1579 (1984).
19. E. F. Hawkins. Solid scintillators for receptor assays: An environmentally safe alternative to liquid scintillation counting. *J. Receptor Res.*, in press (1990).
20. R. M. Burch and J. Axelrod. Dissociation of bradykinin-induced prostaglandin formation from phosphatidylinositol turnover in Swiss 3T3 fibroblasts. *Proc. Natl. Acad. Sci. USA* 84:6374–6378 (1987).
21. B. B. Aggarwal, W. J. Kohr, and P. E. Hass. Human tumor necrosis factor: Production, purification, and characterization. *J. Biol. Chem.* 260:2345–2354 (1985).
22. P. C. Wong, A. T. Chiu, W. A. Price, M. J. M. C. Thoolen, O. J. Carini, A. L. Johnson, R. I. Taber, and P. B. M. W. M. Timmermans. Nonpeptide angiotensin II receptor antagonists. I. Pharmacological characterization of 2-*n*-butyl-4-chloro-1-(2-chlorobenzyl)imidazole-5-acetic acid, sodium salt (5-8307). *J. Pharmacol. Exp. Ther.* 247:1–7 (1988).
23. G. W. Adamson and J. A. Bush. A comparison of the performance of some similarity and dissimilarity measures in the automatic classification of chemical structures. *J. Chem. Inf. Comp. Sci.* 15:55–58 (1975).
24. Y. C. Martin, E. B. Danaher, C. S. May, D. Weininger, and J. H. Van Drie. *J. Comp. Aid. Mol. Des.*, in press.
25. R. P. Sheridan, and R. Venkataraghavan. Designing novel nicotinic agonists by searching a database of molecular shapes. *J. Comp. Aided Mol. des.* 1:243–256 (1987).
26. J. H. Van Drie, D. Weininger, and Y. C. Martin. ALADDIN: an integrated tool for computer-assisted molecular design and pharmacophore recognition from geometric, steric, and substructure searching of three-dimensional molecular structures. *J. Comp. Aided Mol. Des.* 3:225–251 (1987).

Mimicry of erythropoietin by a nonpeptide molecule

Sajjad A. Qureshi*, Ronald M. Kim, Zenon Konteatis, Dawn E. Biazzo, Haideh Motamedi, Robert Rodrigues, Judith A. Boice, Jimmy R. Calaycay, Maria A. Bednarek, Patrick Griffin, Ying-Duo Gao, Kevin Chapman, and David F. Mark†

Merck Research Laboratories, Rahway, NJ 07065

Edited by Donald Metcalf, The Walter and Eliza Hall Institute of Medical Research, Melbourne, Australia, and approved August 10, 1999 (received for review March 30, 1999)

Erythropoietin (EPO) controls the proliferation and differentiation of erythroid progenitor cells into red blood cells. EPO induces these effects by dimerization of the EPO receptors (EPOR) present on these cells. To discover nonpeptide molecules capable of mimicking the effects of EPO, we identified a small molecule capable of binding to one chain of EPOR and used it to synthesize molecules capable of inducing dimerization of the EPOR. We first identified compound 1 (*N*-3-[2-(4-biphenyl)-6-chloro-5-methyl]indolyl-acetyl-L-lysine methyl ester) by screening the in-house chemical collection for inhibitors of EPO binding to human EPOR and then prepared compound 5, which contains eight copies of compound 1 held together by a central core. Although both compounds inhibited EPO binding of EPOR, only compound 5 induced dimerization of soluble EPOR. Binding of EPO to its receptor in cells results in activation of many intracellular signaling molecules, including transcription factors like signal transducer and activator of transcription (STAT) proteins, leading to growth and differentiation of these cells. Consistent with its ability to induce dimerization of EPOR in solution, compound 5 exhibited much of the same biological activities as EPO, such as (*i*) the activation of a STAT-dependent luciferase reporter gene in BAF3 cells expressing human EPOR, (*ii*) supporting the proliferation of several tumor cell lines expressing the human or mouse EPOR, and (*iii*) the *in vitro* differentiation of human progenitor cells into colonies of erythrocytic lineage. These data demonstrate that a nonpeptide molecule is capable of inducing EPOR dimerization and mimicking the biological activities of EPO.

Erythropoietin (EPO) is essential for the maintenance of red blood cells in humans as well as in various animal models (1), reviewed in ref. 2. In humans, the kidney is the primary source of EPO synthesis, whereas other organs such as the liver and brain produce small but significant amounts in adults (3–6). Deficits in EPO production result in anemia in humans and in animal models. In humans, the most prevalent form of anemia is associated with kidney failure (7). At present, the only treatment for this form of anemia is administration of recombinant EPO via subcutaneous or intravenous injection (8–10). The use of recombinant EPO has significantly improved the quality of life of these patients; however, this treatment requires repeated administration of recombinant protein, which is both inconvenient and expensive.

EPO induces its biological effects after binding to a cell-surface receptor (EPOR). Binding of EPO to EPOR results in dimerization of these receptors, as is the case for many other growth factor and cytokine receptors (11, 12). Apparently dimerization of EPOR is all that is required to trigger the biological responses associated with EPO. A constitutively active (hormone-independent) EPOR was first isolated after retroviral transduction (13). The activation of this receptor mapped to an arginine-to-cysteine mutation at position 129 in the human EPOR. The mutant receptor forms disulfide-linked homodimers in the absence of EPO (14). After this example, more constitutively active EPORs have been created by introducing a cysteine residue in parts of the putative EPOR dimerization interphase (15, 16). These mutant receptors, when introduced into growth factor-dependent BAF3 cells, converted them into growth factor-independent cells. Similarly, a bivalent monoclonal antibody directed toward the extracellular domain of the EPOR promotes dimerization of EPOR and mimics EPO activities (17). Moreover, recently a 20-aa peptide, EPO mimetic peptide-1 (EMP-1), has been shown to dimerize the EPOR in solution as well as on the cell surface (18, 19). This peptide exhibits EPO-like activities both *in vitro* and *in vivo* (18).

The crystal structures of EPO/EPOR and EMP-1/EPOR complexes have been solved and reveal a different configuration of the EPOR dimer in each of the complexes (19, 20). On the basis of the three-dimensional structure of EPOR observed in these crystals, the mutations described above are in a region of the exoplasmic domain that is too far away for disulfide bond formation to occur between the two EPOR molecules. Therefore, it is unlikely that the covalently held EPOR dimers induced by mutations in the exoplasmic domain (as described above) will have a configuration similar to that of EPO- or EMP-1-induced EPOR dimers. A similar conclusion may be drawn for the monoclonal antibody-induced EPOR dimers. All these data suggest that, although the dimerization of the EPOR is important, the conformation of EPOR in the dimer complex is quite flexible. This also suggests that other molecules capable of dimerizing the EPOR may be able to act as EPO mimetics as well.

We are interested in developing small-molecule EPO mimetics for the treatment of anemia. On the basis of the characteristics of EPOR-dimerizing entities described above, it appears that such a molecule must have functional groups capable of interacting with at least two receptor chains. One possible way to obtain such a molecule is to first identify a compound that can interact with one chain of the EPOR and then ligate it in such a way that it can now interact with both chains of the receptor. In this paper, we report on the identification of such a molecule, which we initially identified as an EPOR antagonist. This molecule, when presented as an oligomer, is converted into an EPOR agonist, recapitulating some of the biological activities associated with EPO.

Materials and Methods

Synthesis of Compounds and EMP-1. Compound **1** was synthesized by using Starburst polyamidoamino-octa-4-hydroxymethylbenzamide (**2**) as a soluble support, as shown in Fig. 1. To a stirring solution of support **2** (0.05 mmol, 105 mg) (21) and Fmoc-Lys(Boc)-OH (0.8 mmol, 364 mg, 2 eq per handle) in 3 ml of *N,N*-dimethylformamide (DMF) was added catalytic 4-dimeth-

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: EPO, erythropoietin; ^{125}I-EPO, (3-[^{125}I] *iodotyrosyl*) EPO; EPOR, EPO receptor; rEBP, recombinant EPO-binding protein; EMP-1, EPO mimetic peptide-1; STAT, signal transducer and activator of transcription; JAK, janus kinase; CFU, colony-forming unit; GM-CSF, granulocyte–macrophage colony stimulating factor; G-CSF, granulocyte colony stimulating factor; ^{33}P-rEBP, radiolabeled form of rEBP; cpm, counts per minute; DMF, *N,N*-dimethylformamide; SEC, size exclusion chromatography.

*To whom reprint requests should be addressed at: Merck Research Laboratories, P.O. Box 2000, MS RY80Y-310, Rahway, NJ 07065. E-mail: sajjad_qureshi@merck.com.

†Present Address: Hoffmann–La Roche, Inc., 340 Kingsland Street, Nutley, NJ 07110.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

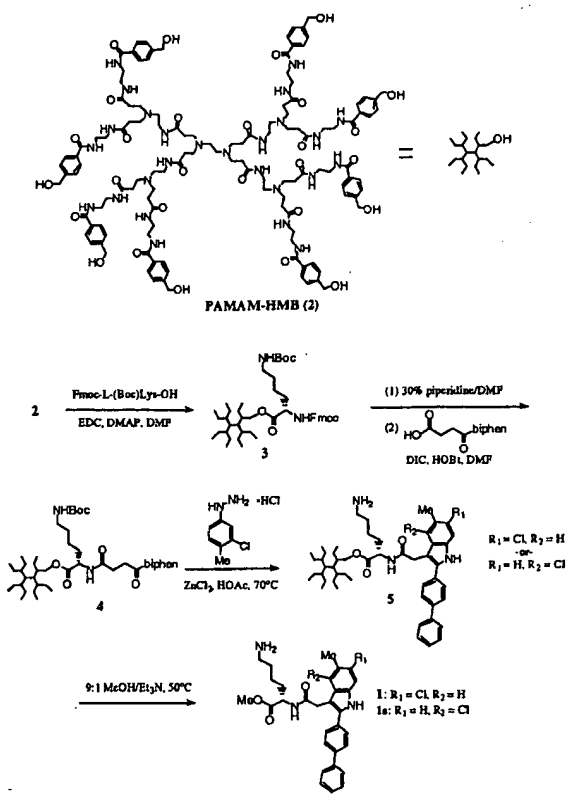


Fig. 1. Scheme 1 for the synthesis of compounds **1** and **5**.

ylaminopyridine (3–5 mg), followed by 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (0.8 mmol, 154 mg). The reaction was stirred overnight at ambient temperature. Dendrimer **3** was isolated by size exclusion chromatography (SEC) on Sephadex LH-20 (Amersham Pharmacia) eluting with DMF, followed by SEC on Bio-Rad Biobeads S-X3 eluting with CH_2Cl_2. Removal of solvent *in vacuo* afforded 204 mg of **3** as a beige solid. Compound **3** (0.016 mmol, 100 mg) was treated with 3 ml of 30% piperidine in DMF for 15 min and isolated by SEC on Sephadex LH-20. The resulting octa-amine was dried *in vacuo*, and residue was dissolved in 3 ml of DMF containing 3-(4-biphenylbenzoyl)propionic acid (0.3 mmol, 76 mg, 2.3 eq) and 1-hydroxybenzotrazole hydrate (0.3 mmol, 41 mg). To the solution was added diisopropylcarbodiimide (0.3 mmol, 54 μL). The mixture was allowed to stand overnight, and the product was isolated by SEC on Biobeads S-X3 eluting with CH_2Cl_2. Removal of solvent afforded 114 mg of dendrimer **4** as a beige solid. Dendrimer **4** (84 mg) and 3-chloro-4-tolylhydrazinium chloride (1.5 mmol, 290 mg) were dissolved in 3 ml of glacial acetic acid containing zinc chloride (1.5 mmol, 204 mg) and anisole (0.6 mmol, 60 mg, 60 μL), and the slurry was heated overnight at 70°C. The solvent was removed *in vacuo*, and the residue was dissolved in DMF and purified by SEC on Sephadex LH-20. Removal of solvent afforded 68 mg of **5** as a tan solid. ^{1}H NMR of **5** revealed the removal of the Boc-protecting group during the reaction. Dendrimer **5** (40 mg) was taken up in 2 ml of 9:1 MeOH/Et_3N and heated at 50°C for 22 hr. The mixture was rotary evaporated to a tan film, which was taken up in MeCN. The insoluble dendrimer was filtered, and the light-colored filtrate was rotary evaporated to afford 17 mg of methyl esters **1** and **1b** as the major products, in addition to the corresponding cyclized lactams. Separation of the two regioisomers by reverse-phase chromatography (C-8, H_2O/MeCN gradient containing 0.15% trifluoroacetic acid) afforded 1.7 mg of *N*-3-[2-(4-biphenyl)-6-chloro-5-methyl]indolyl-acetyl-L-lysine methyl ester (**1**) and 2.7 mg of *N*-3-[2-(4-biphenyl)-4-chloro-5-methyl]indolyl-acetyl-L-lysine methyl ester (**1a**) as the trifluoroacetate salts. **1** ^{1}H NMR (400 MHz, CD_3OD): 7.76 (d, 2H), 7.70 (d, 2H), 7.68 (d, 2H), 7.47 (t, 2H), 7.38 (t, 1H),7.27 (d, 1H), 7.06 (d, 1H), 4.58 (m, 1H), 4.04 (d, 2H), 3.72 (s, ^{3}H), 2.87 (t, 2H), 2.43 (s, ^{3}H), 1.85–1.98 (m, 1H), 1.57–1.70 (overlapping m, ^{3}H), 1.43 (m, 2H). **1a** ^{1}H NMR (400 MHz, CD_3OD): d7.78 (d, 2H), 7.75 (d, 2H), 7.68 (d, 2H), 7.48 (s, 1H), 7.46 (t, 2H), 7.41 (s, 1H), 7.36 (t, 1H), 4.52 (m, 1H), 3.86 (s, 2H), 3.70 (s, ^{3}H), 2.81 (m, 2H), 2.45 (s, ^{3}H), 1.90 (m, 1H), 1.73 (m, 1H), 1.63 (m, 2H), 1.40 (m, 2H). Dried powders were stored at ambient temperatures and used for preparation of stock solutions in 100% DMSO. Stock solutions prepared in this way were stored at −20°C until use and were used within a few months, during which no loss of activity was observed as determined by activities of these compounds in the EPOR-binding and luciferase-reporter gene assays described here. EMP-1, GGTYcyclo(CHFGPLTWVC)KPQGG-amide, was prepared by the solid-phase method (22) on a 431A Applied Biosystems peptide synthesizer and dissolved in 100% DMSO.

Cell Lines. Murine BAF3 cells (kindly provided by Alan D'Andrea, Harvard Medical School, Boston), DA3 cells (kindly provided by Jim Ihle, St. Jude Children's Research Hospital, Memphis, TN), and cell lines derived from these were maintained in RPMI-1640 supplemented with 10% FBS, antibiotic, L-glutamine (all from GIBCO/BRL), and 5 ng/ml mIL-3 (R & D systems) at 37°C in a humidified incubator. Human erythroleukemia cell line F36 (obtained from Riken Cell Bank, Tsukuba Science City, Japan) and TF-1 (obtained from K. Kitamura, Tokyo University, Tokyo) (23, 24) were maintained in the above-mentioned media supplemented with hIL-3 (5 ng/ml).

Plasmids. pET15b/hEPOR-ECD expresses the extracellular domain of the human EPOR in *Escherichia coli*, and RcCMV/EPOR expresses human EPOR in mammalian cells and will be described elsewhere (H.M., unpublished data). pAH4-LUC contains six copies of the signal transducer and activator of transcription (STAT)-binding site from the interferon regulatory factor 1 gene cloned upstream of the herpes simplex virus–thymidine kinase minimal promoter and the luciferase gene (25). A gene cassette conferring resistance to the antibiotic Zeocine (Invitrogen) was introduced at a *Sal*I site in the pAH4LUC to yield pAH4LUCZeo.

Expression and Purification of EPO-Binding Protein in *E. coli*. *E. coli* strain BL21 DE3 (Novagen) containing pET15b/hEPOR-ECD was used for expression of the extracellular domain of the human EPOR. The resulting recombinant protein was processed to yield the recombinant EPO-binding protein (rEBP) as described previously (26). An EPO-dependent luciferase assay was used to further characterize the rEBP. The purified rEBP inhibited EPO-induced luciferase activity in this assay with an IC_{50} of 5–10 nM (S.Q. and R.R., unpublished data).

EPOR-Binding Assay. One microgram of the purified His-tagged-rEBP in 70 μl of 1× PBS [1× PBS contains 137 mM NaCl, 2.68 mM KCl, 1.46 mM KH_2PO_4, 15.8 mM sodium phosphate, pH 7.2 (GIBCO/BRL)] containing 0.05% sodium azide was dispensed into 96-well high-binding microtiter plates (Costar 3922) and allowed to bind at 4°C for at least 24 hr. All unbound rEBP was removed by washing with 1× PBS containing 0.05% Tween 20 (Sigma), and nonspecific binding sites were blocked by incubating with 1% BSA (Pierce) in 1× PBS and 0.05% Tween 20. The plate-immobilized rEBP bound ([3-[^{125}I]iodotyrosyl]) EPO (^{125}I-EPO) with a K_d of ≈5 nM. All competition binding experiments were performed in a volume of 50 μl containing 5 nM ^{125}I-EPO (Amersham, specific activity 300–900 Ci/mmol) in 1× PBS/

0.5% BSA/0.05% Tween 20/0.01% sodium azide/5% DMSO +/− compounds. To set up these assays, each compound was diluted into the assay buffer (from the stocks that have been prepared in 100% DMSO) to yield the appropriate compound concentration and added to each well of a 96-well plate. These reactions were incubated at 4°C for 16–20 hr to allow equilibrium binding, and unbound ^{125}I-EPO was removed by five rapid washes with cold 1× PBS containing 0.5% BSA. Bound radioactivity was counted, after addition of 100 μl of Microscint (Packard)/well, in a Topcount scintillation counter (Packard). Under these conditions, unlabeled EPO inhibited binding of ^{125}I-EPO with an IC_{50} of ≈5 nM.

EPOR Dimerization Assay. A detailed description of this assay will be published elsewhere (D.B. and S.Q., unpublished work). Briefly, rEBP containing a protein kinase A substrate site was produced in *E. coli* as described above. This protein was radiolabeled by using protein kinase A and ^{33}P-γATP to high specific activity. The soluble ^{33}P- rEBP was allowed to interact with the plate-bound rEBP as described for the receptor-binding assay. Each reaction was performed in a 100 μl volume containing 1- to 2-nM-labeled rEBP in 50 mM Hepes (pH 7.2)/5 mM Mg_2Cl/5 mM Ca_2Cl/0.05% sodium azide/1% BSA/5% DMSO/+/−compounds) in 96-well microtiter plates in the presence or absence of 250 ng/well immobilized rEBP. After incubation for 16–20 hr in the presence or absence of compounds, plates were washed, and the amount of radioactivity was determined as described in the EPOR-binding assay.

Establishment of Luciferase Reporter Cell Lines in BAF3 Cells. The pAH4LUCZeo DNA was linearized at a *PvuI* site and transfected together with pHOOK3 (Invitrogen) into the BAF3 cells by electroporation. Forty-eight hours after transfection, the cells transfected with pHOOK3 were separated from the untransfected cells by using magnetic beads coated with phOx per manufacturer's suggested protocol (Invitrogen). The isolated cells were plated at one cell per well in 96-well microtiter plates and allowed to grow in the presence of Zeocine (200 μg/ml) for selection of stable transfectants. Zeocine-resistant cells were tested in a luciferase assay after induction with murine IL-3 to establish inducibility of the luciferase reporter gene in these cells. Clones responsive to murine IL-3 (BAF3/LUC cells) were further expanded and used for introduction of human EPOR. For introduction of human EPO, BAF3/LUC cells maintained in Zeocine were electroporated in the presence of RcCMV/EPOR and selected in the presence of 200 μg/ml Geneticin (GIBCO/BRL) for 2 wk. The cultures were further selected in EPO (1 unit/ml) to obtain a pool of BAF3/LUC cells expressing the functional EPOR receptor.

Luciferase Assay. Fifty thousand cells in 200 μl of RPMI 1640 containing 10% FBS were plated into each well of 96-well plates (Falcon). After 16–20 hr of incubation with cytokines or compounds, the luciferase activity was determined after addition of an equal volume of LucLite Reagent (Packard) and luminescence measured in a MLX Microtiterplate Luminometer (Dynatech).

Mitogenic Assay. Mouse BAF3 cells expressing human EPOR, DA3 cells expressing human and mouse EPOR, parental BAF3 and DA3 cells, and human erythroleukemia cell lines F36 and TF-1 were plated at a density of 10,000–20,000 cells per well in 96-well microtiter plates in RPMI 1640 containing 10% FBS. After 40 hr of incubation with cytokines or compounds, ^{3}H-thymidine (4 μCi/ml) was added to each well and allowed to incorporate into DNA for 4 hr. The cells were harvested onto a LKB Filtermat B by using a Tomtec Cell Harvester Mach II (Tomtec, Orange, CT). Filtermats were dried, sealed in counting bags with 23-ml scintillation mixture, and counted on a Wallac (Gaithersburg, MD) 1205 Betaplate counter.

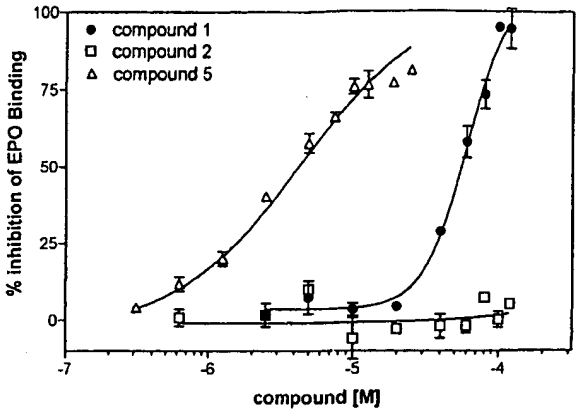


Fig. 2. Inhibition of ^{125}I-EPO binding to rEBP. Dose-response curve for inhibition of ^{125}I-EPO binding by compounds **1**, **2**, and **5**. Competition binding experiments were performed in a volume of 50 μl containing 5 nM ^{125}I-EPO in the presence or absence of compounds and allowed to proceed for at least 16 hr at 4°C, as described in *Materials and Methods*. Nonspecific EPO binding was determined by performing the experiments in the presence of 1 μM unlabeled EPO and was found to be less than 10% (under these conditions, we routinely obtained a total of 3,000–5,000 counts per minute (cpm)/well with no compound and 200–400 cpm/well in the presence of 1 μM cold EPO). The data are expressed as a percentage of the EPO (specific cpm) bound in the presence of 5% DMSO but in the absence of any compound (which is considered as 100%). Each data point was analyzed in triplicate and is a mean (+/− SEM) of three independent experiments.

Hematopoietic Colony Assays. Colony assays were set up by using human $CD34^{+}$ mononuclear cells isolated from healthy donors. One thousand to two thousand cells were mixed with 1 ml of methylcellulose culture medium containing FBS (StemCell Technologies, Vancouver), 0.4% DMSO, a mixture of growth factors containing granulocyte–macrophage colony stimulating factor (GM-CSF), IL-3, granulocyte colony stimulating factor (G-CSF), and stem cell factor (HC-4535-StemCell Technologies, Vancouver), and EPO (Epoetin Alfa, Amgen Biologicals) or compounds. After thorough mixing, the suspension was plated in 35-mm gridded plates (Nunc) and incubated at 37°C in a humidity controlled CO_2 incubator. Colony-forming units (CFU)–erythroid containing >50 hemoglobinized cells and myeloid colonies (CFU-GM) containing >50 cells were counted on day 12–14. Mixed colonies, containing both erythroid and myeloid cells, were counted on day 14–16. All assays were performed at Poietic Technologies, Gaithersburg, MD.

Results and Discussion

Compound 5, an Antagonist of EPO-Binding to EPOR, also Induces Dimerization of the EPOR. To test the hypothesis that a multimeric form of an EPOR antagonist can function as an EPO mimetic to dimerize and activate the EPOR, we screened the in-house chemical collection for inhibitors of EPO binding to the EPOR using the extracellular domain of the human EPOR as rEBP. This screening resulted in identification of *N*-3-[2-(4-biphenyl)-6-chloro-5-methyl]indolyl-acetyl-L-lysine methyl ester (compound **1**). Compound **1** inhibited binding of ^{125}I-EPO to the rEBP with an IC_{50} of 59.5 (+/− 1.1) μM (Fig. 2). Compound **5**, a precursor in the synthesis of compound **1**, consists of eight copies of compound **1** attached to a polyamidoamino-octa-4-hydroxymethylbenzamide support via a chemical linker (compound **2**). When tested in the EPOR binding assay, compound **5** also inhibited binding of ^{125}I-EPO to the rEBP with an IC_{50} of

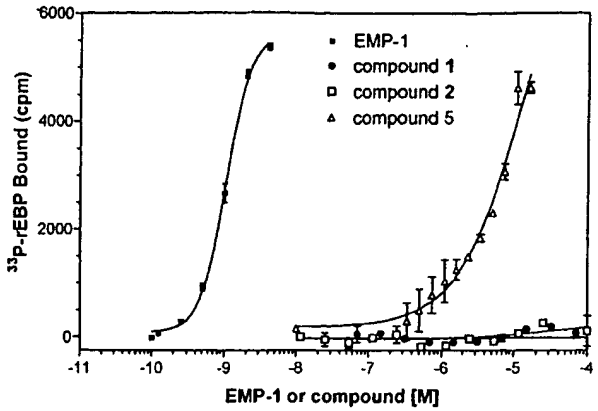


Fig. 3. Dimerization of soluble EPOR. ^{33}P-rEBP was incubated with or without plate-immobilized rEBP (as in Fig. 2), in the absence or presence of compounds **1**, **2**, or **5** or EMP-1 for 16 hr at 4°C, as described in *Materials and Methods*. Data represent specific cpm (cpm with plate-immobilized EPOR minus cpm without plate-immobilized EPOR) observed with each of the compounds. Data are a mean of two experiments (+/− SEM) performed in triplicate. Under these conditions, mean cpm observed in wells with no compounds were 147 +/− 25).

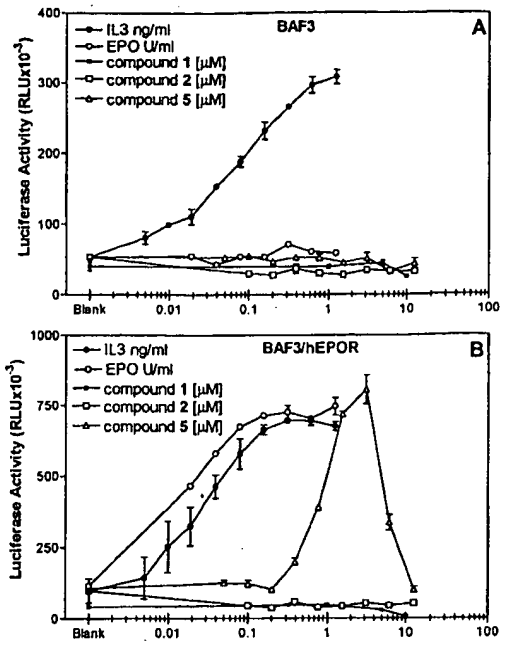


Fig. 4. Induction of luciferase activity in BAF3/LUC and BAF3/LUC/EPOR cells. BAF3/LUC or BAF3/LUC/EPOR cells containing a stably integrated luciferase gene under the control of a STAT-binding element were treated with varying amounts of EPO, IL-3, or compounds for 16 hr. The luciferase activity was determined as described in *Materials and Methods* and is expressed in relative light units (RLU). (*A*) Luciferase activity in BAF3/LUC cells. (*B*) Luciferase activity in BAF3/LUC/EPOR cells, expressing human EPOR. Data are a mean (+/− SEM) of two to three independent experiments performed in triplicate. Value of *x* axis reflects the concentration of EPO, IL-3, or compounds in unit/ml, ng/ml, and μM, respectively. All assays were performed in 1% DMSO. Values shown above blank on graphs refer to luciferase activity observed in the untreated cultures in the presence of 1% DMSO.

4.4 +/− 1.9 μM (Fig. 2). Compound **2**, the unconjugated dendrimer used to crosslink compound **1**, did not affect binding of ^{125}I-EPO to the rEBP in this assay. The ability of compound **5** to inhibit binding of ^{125}I-EPO to the rEBP suggests that compound **1**, even when attached to an inactive central core, is able to interact with rEBP.

The biological activation of EPOR requires dimerization of EPOR on binding of EPO to two independent chains of EPOR. Thus, one of several ways for a small molecule to activate the EPOR is to induce dimerization of the EPOR. To determine whether compound **1** or **5** can dimerize EPOR in solution, we tested these compounds in an EPOR dimerization assay. This assay examined the interaction of a radiolabeled form of rEBP (^{33}P-rEBP) in solution with a plate-immobilized rEBP. In the absence of any dimerizing agent in the assay, the labeled receptor fails to attach to the plate and is easily washed off during the washing step (at the end of the assay). However, in the presence of a compound that can interact with at least two chains of EPOR, the radiolabeled chain would be expected to remain bound to the plate, resulting in increased retention of ^{33}P-rEBP in the assay. EMP-1, a 20-aa peptide, which has been shown to induce dimerization of this form of EPOR in solution (18), produces a dose-dependent increase in the retention of ^{33}P-rEBP (Fig. 3). When tested in this assay, compound **5** produced a dose-dependent increase (EC_{50} 15.9 +/− 3.3 μM) in retention of ^{33}P-rEBP, whereas compounds **1** and **2** both failed to produce any such increases (Fig. 3). These data suggest that even though compounds **1** and **5** both compete for the binding of EPO to the rEBP, only compound **5** is able to interact with more than one chain of EPOR.

Compound 5 Induces Cellular Transcription and Proliferation in Cells Expressing EPOR. Binding of EPO to its cell-surface receptors results in the activation of cellular signaling pathways, including activation of janus kinase (JAK)2 and STAT5 (27–29). STAT5 is a transcription factor that translocates to the nucleus on activation and induces gene transcription (30). Luciferase reporter gene constructs containing synthetic promoters with STAT-binding sites have been used as markers for the activation of STATs in cells (31–33). To determine whether compound **5** could interact with the EPOR expressed in mammalian cells, induce receptor oligomerization as seen *in vitro* (Fig. 3), and activate the JAK-STAT signaling pathway, we tested this compound on BAF3/LUC cells. The BAF3/LUC cells have a stably integrated luciferase gene under the control of a STAT-binding site from the interferon regulatory factor 1 gene. This STAT-binding site has been shown to function in response to activation signals involving STATs 1–5 (25, 34–36). Because BAF3 cells do not normally express EPORs, they do not manifest any biological response when treated with EPO; however, they can acquire EPO responsiveness once the EPOR is expressed ectopically (37, 38). However, these cells do express IL-3 receptors and in response to IL-3 activate the JAK2 and STAT5 signaling pathway (39–41). As shown in Fig. 4, treatment with IL-3 increases the luciferase activity in both BAF3/LUC and BAF3/LUC/EPOR cells, demonstrating that these cells are competent in responding to receptor-dependent activation of Stat5. As expected, EPO-dependent increases in the luciferase activity were observed in the EPO-treated BAF3/LUC/EPOR cells but not in EPO-treated BAF3/LUC cells. When tested in these cells, compound **5** induced increases in the luciferase activity only in the BAF3/LUC/hEPOR cells with an EC_{50} of 1.25 μM. No such increases were observed in BAF3/LUC cells (Fig. 4). Moreover, neither compound **1** nor compound **2** induced any increases in the luciferase activity in either of these two cell lines. The decrease in the luciferase activity observed at higher concentrations of compound **5** may be caused by toxicity or by the engagement of EPOR in a 1:1 (EPOR/compound) complex, thus preventing the formation of EPOR dimers needed for the activation of the JAK-STAT pathway. Formation of this 1:1 receptor/ligand complex at high ligand concentration has been

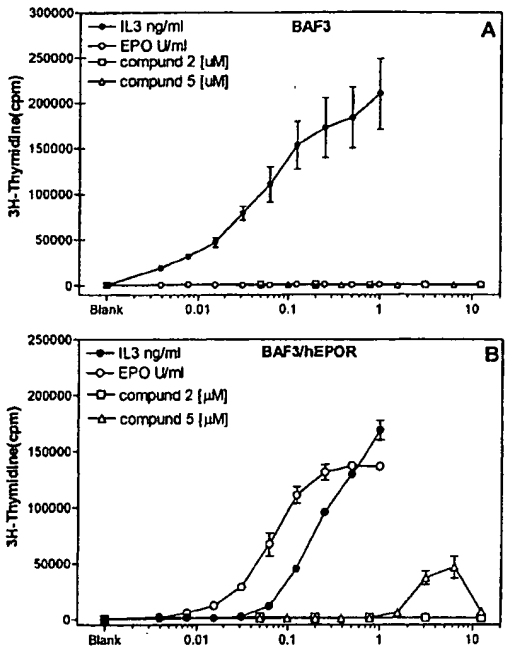


Fig. 5. Mitogenic responses in BAF3 and BAF3/hEPOR cells. BAF3 or BAF3/hEPOR cells were treated with different amounts of EPO, IL-3, or compounds for 48 hr. ^{3}H-thymidine incorporation, as an indicator of cell proliferation, was measured after addition of 4 μCi/ml of ^{3}H-thymidine during the last 4 hr of incubation, as described in *Materials and Methods*. (*A*) Amount of ^{3}H-thymidine incorporated in BAF3 cells after different treatments. (*B*) Effect of same in BAF3 cells expressing human EPOR. Data represent mean (+/− SEM) of two to three experiments where each determination was made in triplicate. Value of *x* axis reflects the concentration of EPO, IL-3, or compounds in units/ml, ng/ml, and μM, respectively. All assays were performed in 1% DMSO. Values shown above blank on graphs refer to level of radioactivity incorporated in the untreated cultures in the presence of 1% DMSO.

shown to inhibit the activation of growth hormone and EPORs by their respective ligands (17, 42).

BAF3 cells are known to proliferate in the presence of IL-3 and a number of other cytokines, including EPO, provided that the receptors for these cytokines are expressed in these cells. The increased proliferation results in increased incorporation of nucleotides, such as a thymidine, into DNA, which can be measured by using ^{3}H-thymidine. Much like what was observed for the luciferase reporter gene activation, IL-3 induced increases in the ^{3}H-thymidine incorporation in both BAF3 and BAF3/hEPOR cells, but EPO induced these increases in BAF3/hEPOR cells only (Fig. 5). When tested in these cells, compound **5** induced ^{3}H-thymidine incorporation in BAF3/hEPOR cells but not in parental BAF3 cells, similar to the response to EPO. However, compound **2** failed to show any activity in the assay (because of lack of activity in the EPOR dimerization and luciferase reporter gene assay, compound **1** was not evaluated in this or any subsequent assays), and again compound **5** showed a decreased activity at higher concentrations, which could be caused by nonspecific cellular toxicity or unproductive engagement of receptors (as suggested above). BAF3 cells require the presence of a cytokine such as IL-3 even for survival in cell culture, thus making it difficult to distinguish between these two possibilities. However, this phenomenon is not limited to BAF3 cells, because we have observed similar inhibition at high concentrations in DA3/hEPOR cells, another murine cell line, as well as in F36E and TF-1 cells, two cell lines of human origin known to respond to EPO via the endogenous EPOR (23, 24). Moreover, we have observed a compound **5**-dependent inhibition of IL-3-induced thymidine incorporation in BAF3 cells lacking the EPOR (S.Q. and R.R., unpublished observation), which supports the notion that the toxicity of compound **5** is nonspecific and independent of the EPOR. More experiments will be needed to sort out this phenomenon. Regardless of the complexity in the interpretation of the data at higher doses of compound **5**, these data still demonstrate that compound **5** interacts with the EPOR expressed on the cell surface, resulting in induction of EPO-like activities.

Table 1. Induction of CFU-erythroid colonies by EPO and compound 5

Treatment	Dose	Avg. no. colonies per treatment: Erythroid colonies	Avg. no. colonies per treatment: GM colonies
None	—	0 ± 0	86 ± 13
DMSO	0.4%	0 ± 0	74 ± 2
EPO	100 mU/ml	66 ± 1	87 ± 7
DMSO + EPO	100 mU/ml	55 ± 2	73 ± 5
	30 mU/ml	48 ± 3	73 ± 5
	10 mU/ml	20 ± 7	72 ± 13
	3 mU/ml	2 ± 1	65 ± 5
Compound **2**	2000 nM	0 ± 0	65 ± 15
	500 nM	0 ± 0	66 ± 11
	125 nM	0 ± 0	72 ± 11
Compound **5**	2000 nM	47 ± 2	65 ± 1
	1000 nM	43 ± 1	76 ± 1
	500 nM	30 ± 2	75 ± 5
	250 nM	12 ± 2	71 ± 1
	125 nM	3 ± 1	75 ± 1

Human CD34$^+$ progenitor cells (1.5 × 10^3/plate) were plated in semisolid media to allow growth of erythroid colonies and GM colonies. Colonies of each type were counted after 12–15 days of seeding under each condition. Data are mean (SEM) of two independent experiments, where each determination was performed in triplicate.

Compound 5 Induces Differentiation of Human Progenitor Cells into Cells of Erythrocytic Lineage. EPO induces differentiation of hematopoietic progenitor cells into mature erythrocytes *in vivo*. However, in tissue culture conditions, only a part of this process is recapitulated. Progenitor cells isolated from peripheral blood or bone marrow when cultured *in vitro* in semisolid media can grow into morphologically distinct colonies (each colony being a product of single progenitor cell), depending on which growth factor is present in the media. For example, it is well known that EPO, in the presence of stem cell factor and IL-3, induces the formation of colonies known as CFU-E (*C*olony *F*orming *U*nit–*E*rythroid), which are easily identified microscopically because of the hemoglobinization of cells in the colony. Cells isolated from human peripheral blood by using an anti-CD34 antibody (CD34$^+$ cells) are more advanced progenitors capable of differentiation into erythrocytes in the presence of EPO and into granulocytes and macrophages in the presence of G-CSF and GM-CSF (43, 44). We used these cells to examine the ability of compound **5** to replace EPO in promoting differentiation of these progenitor cells into erythroid colonies *in vitro*. Cells plated in the presence of growth factors (stem cell factor, IL-3, GM-CSF, and G-CSF) alone did not develop into erythroid colonies, even though these were able to form colonies of GM progenitor cells (CFU-GM) indicating the viability of these isolated cells (Table 1). However, addition of EPO to the mixture of growth factors induced formation of erythroid colonies in a dose-dependent manner. Under the same conditions, compound **5** also induced erythroid colonies, whereas compound **2** did not

induce erythroid colonies and did not inhibit the formation of GM colonies. Moreover, no morphological differences could be observed between the EPO-induced erythroid colonies or the compound-induced erythroid colonies. These data confirm that compound **5** is an EPO mimetic capable of inducing the proliferation and differentiation of human erythroid progenitor cells.

In summary, we have reported on the identification and characterization of an EPOR antagonist, which when presented in a multivalent configuration induced biological responses similar to what is normally observed with EPO. Although the potency of the activities observed with this compound is only a fraction of what is expected of EPO, it does validate the concept that the EPOR, and by extension most cytokine receptors, can be ligated together in an active conformation by a nonpeptidyl molecule. The only requirement is that the small molecule must be able to interact with both chains of the receptor. This paper also lays out a basic strategy for identifying cytokine mimetics by converting an antagonist into an agonist.

We thank Drs. Alan D'Andrea, Jim Ihle, and K. Kitamura and the Rieken Cell Bank for cell lines used in these studies, and Poietic Technologies, Inc. for performing the hematopoietic colony assays. We also thank Steve Parent, Kristine Prendergast, and Bruce Bush for helpful discussions, Bruce Bush, David E. Moller, and Russell B. Lingham for comments on the manuscript, and Roy Smith and Bennett Shapiro for their encouragement and support.

1. Koury, M. J. & Bondurt, M. C. (1988) *J. Cell. Physiol.* **137,** 65–74.
2. Krantz, S. (1991) *Blood* **77,** 419–434.
3. Koury, S. T., Bondurant, M. C. & Koury, M. J. (1988) *Blood* **71,** 524–527.
4. Lacombe, C. D. & De Silva, J. L. (1988) *J. Clin. Invest.* **81,** 621–623.
5. Jelkmann, W. (1992) *Physiol. Rev.* **72,** 449–489.
6. Maxwell, P. H., Ferguson, D. J. P., Nicholls, L. G., Iredale, J. P., Pugh, C. W., Johnson, M. H. & Ratcliffe, P. J. (1997) *Kidney Int.* **51,** 393–401.
7. Valderrabano, F. (1996) *Kidney Int.* **50,** 1373–1391.
8. Spivak, J. L. (1993) *Annu. Rev. Med.* **44,** 243–253.
9. Zhang, F., Gagnon, R. F., Richards, G. K. & Brox, A. G. (1996) *Nephron* **72,** 654–661.
10. Brox, A. G., Zhang, F., Guyda, H. & Gagnon, R. F. (1996) *Kidney Int.* **50,** 937–943.
11. Wells, J. (1994) *Cell Biol.* **6,** 163–173.
12. Heldin, C.-H. (1995) *Cell* **80,** 213–223.
13. Yoshimura, A., Longmore, G. & Lodish, H. (1990) *Nature (London)* **348,** 647–649.
14. Watowich, S., Yoshimura, A., Longmore, G., Hilton, D., Yoshimura, Y. & Lodish, H. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 2140–2144.
15. Watowich, S. S., Hilton, D. J. & Lodish, H. F. (1994) *Mol. Cell. Biol.* **14,** 3535–3549.
16. Longmore, G., Pharr, P. & Lodish, H. (1994) *Mol. Cell. Biol.* **14,** 2266–2277.
17. Schneider, H., Chaovapong, W., Matthews, D. J., Karkaria, C., Cass, R. T., Zhan, H. J., Boyle, M., Lorenzini, T., Elliott, S. G. & Giebel, L. B. (1997) *Blood* **89,** 473–482.
18. Wrighton, N. C., Farrell, F. X., Chang, R., Kashyap, A. K., Barbone, F. P., Mulcahy, L. S., Johnson, D. L., Barrett, R. W., Jolliffe, L. K. & Dower, W. J. (1996) *Science* **273,** 458–463.
19. Livnah, O., Stura, E. A., Johnson, D. L., Middleton, S. A., Mulcahy, L. S., Wrighton, N. C., Dower, W. J., Jolliffe, L. K. & Wilson, I. A. (1996) *Science* **273,** 464–471.
20. Syed, R. S., Reid, S. W., Li, C., Cheetham, J. C., Aoki, K. N., Liu, B., Zhan, H., Osslund, T. D., Chirino, A. J., Zhang, J., *et al.* (1998) *Nature (London)* **395,** 511–516.
21. Kim, R. M., Manna, M., Hutchins, S. M., Griffin, P. R., Yates, N. A., Bernick, A. M. & Chapman, K. T. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 10012–10017.
22. Penninghton, M. W. & Dunn, B. M. (194) in *Methods in Molecular Biology* (Humana, Totowa, NJ), Vol. 35.
23. Chiba, S., Takaku, F., Tange, T., Shibuya, K., Misawa, C., Sasaki, K., Miyagawa, K., Yazaki, Y. & Hirai, H. (1991) *Blood* **78,** 2261–2268.
24. Kitamura, T., Tange, T., Terasawa, T., Chiba, S., Kuwaki, T., Miyagawa, K., Piao, Y.-F., Miyazono, K., Urabe, A. & Takaku, F. (1989) *J. Cell. Physiol.* **140,** 323–334.
25. Rosenblum, C. I., Tota, M., Cully, D., Smith, T., Collum, R., Qureshi, S., Hess, J. F., Phillips, M. S., Hey, P. J., Vongs, A., *et al.* (1996) *Endocrinology* **137,** 5178–5181.
26. Johnson, D., Middleton, S., McMahon, F., Barbone, F., Kroon, D., Tsao, E., Lee, W., Mulcahy, L. & Jolliffe, L. (1996) *Protein Expression Purif.* **7,** 104–113.
27. Gouilleux, F., Pallard, C., Dusanterfourt, I., Wakao, H., Haldosen, L. A., Norstedt, G., Levy, D. & Groner, B. (1995) *EMBO J.* **14,** 2005–2013.
28. Quelle, F. W., Wang, D., Nosaka, T., Thierfelder, W. E., Stravopodis, D., Weinstein, Y. & Ihle, J. N. (1996) *Mol. Cell. Biol* **16,** 1622–1631.
29. Ihle, J. N. (1996) *Philos. Trans. R. Soc. London B* **351,** 159–166.
30. Penta, K. & Sawyer, S. T. (1995) *J. Biol. Chem.* **270,** 31282–31287.
31. Wakao, H., Gouilleux, F. & Groner, B. (1994) *EMBO J.* **13,** 2182–2191.
32. Lamb, P., Seidel, H. M., Haslam, J., Milocco, L., Kessler, L. V., Stein, R. B. & Rosen, J. (1995) *Nucleic Acids Res.* **23,** 3283–3289.
33. Seidel, H. M., Milocco, L. H., Lamb, P., Darnell, J. E., Stein, R. B. & Rosen, J. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 3041–3045.
34. Rothman, P., Kreider, B., Azam, M., Levy, D., Wegenka, U., Eilers, A., Decker, T., Horn, F., Kashleva, H., Ihle, J. & Schindler, C. (1994) *Immunity* **1,** 457–468.
35. Schindler, C. (1995) *Receptor* **5,** 51–62.
36. Darnell, J. E. (1997) *Science* **277,** 1630–1635.
37. Barber, D. L., Mason, J. M., Fukazawa, T., Reedquist, K. A., Druker, B. J., Band, H. & D'Andrea, A. D. (1997) *Blood* **89,** 3166–3174.
38. Pless, M., Norga, K., Carroll, M., Heim, M. H., Dandrea, A. D. & Matheyprevot, B. (1997) *Blood* **89,** 3175–3185.
39. Azam, M., Erdjument-Bromage, H., Kreider, B. L., Xia, M., Quelle, F., Basu, R., Saris, C., Tempst, P., Ihle, J. N. & Schindler, C. (1995) *EMBO J.* **14,** 1402–1411.
40. Yoshimura, A., Ichihara, M., Kinjyo, I., Moriyama, M., Copeland, N. G., Gilbert, D. J., Jenkins, N. A., Hara, T. & Miyajima, A. (1996) *EMBO J.* **15,** 1055–1063.
41. Matsumoto, A., Masuhara, M., Mitsui, K., Yokouchi, M., Ohtsubo, M., Misawa, H., Miyajima, A. & Yoshimura, A. (1997) *Blood* **89,** 3148–3154.
42. Fuh, G., Cunningham, B., Fukunaga, R., Nagata, S., Goeddel, D. & Wells, J. (1992) *Science* **256,** 1677–1680.
43. Testa, U., Fossati, C., Samoggia, P., Masciulli, R., Mariani, G., Hassan, H. J., Sposi, N. M., Guerriero, R., Rosato, V., Gabbianelli, M., *et al.* (1996) *Blood* **88,** 3391–3406.
44. Alcorn, M. J., Richmond, H. L., Pearson, C., Farrel, E., Kyle, B., Dunlop, D. J., Fitzsimons, E., Steward, W. P., Pragnell, I. B. & Franklin, I. M. (1996) *J. Clin. Oncol.* **14,** 1839–1847.

A Small, Nonpeptidyl Mimic of Granulocyte–Colony-Stimulating Factor

Shin-Shay Tian,* Peter Lamb,*† Andrew G. King, Stephen G. Miller, Linda Kessler, Juan I. Luengo, Laurie Averill, Randall K. Johnson, John G. Gleason, Louis M. Pelus, Susan B. Dillon, Jonathan Rosen

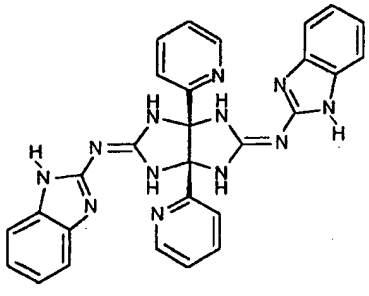
A nonpeptidyl small molecule SB 247464, capable of activating granulocyte–colony-stimulating factor (G-CSF) signal transduction pathways, was identified in a high-throughput assay in cultured cells. Like G-CSF, SB 247464 induced tyrosine phosphorylation of multiple signaling proteins and stimulated primary murine bone marrow cells to form granulocytic colonies in vitro. It also elevated peripheral blood neutrophil counts in mice. The extracellular domain of the murine G-CSF receptor was required for the activity of SB 247464, suggesting that the compound acts by oligomerizing receptor chains. The results indicate that a small molecule can activate a receptor that normally binds a relatively large protein ligand.

Activation of transmembrane receptors for growth factors and cytokines occurs when oligomerization of receptor chains is triggered by binding of a protein ligand to a specific ligand-binding domain on the receptor (*1, 2*). The resultant clustering of tyrosine kinase domains on the cytoplasmic side of the receptor initiates a series of signal transduction events that ultimately alter cellular phenotype. Receptors can also be activated by bivalent receptor antibodies (*2, 3*) and by dimeric peptides that interact with the ligand binding domain (*4*), which also induce receptor oligomerization. Activation of receptors by small, nonpeptidyl molecules amenable to chemical synthesis would make possible the development of orally available growth factor and cytokine mimics.

The protein hormone granulocyte–colony-stimulating factor (G-CSF) has a primary role in the production and activation of cells of the granulocytic lineage (*5*). Recombinant G-CSF is used to treat a variety of congenital and iatrogenic human neutropenias (*6*). Binding of G-CSF to its receptor triggers receptor homodimerization, which leads to activation of two members of the JAK family of protein tyrosine kinases, JAK1 and JAK2, which associate with the cytoplasmic domain of the receptor (*7, 8*). The activated JAKs phosphorylate tyrosine residues on the cytoplasmic face of the receptor, which then serve as the binding sites for signaling proteins. The JAKs are then presumed to phosphorylate the receptor-associated proteins, among which are the STATs (signal transducers and activators of transcription). After phosphorylation on tyrosine, the STATs dimerize, translocate to the nucleus, and bind to specific DNA sequences in the promoters of responsive genes, thereby regulating transcription (*7, 9, 10*).

We developed a high-throughput, cell-based screen to detect compounds that activate the G-CSF receptor. The screen relies on a reporter gene driven by a synthetic STAT-responsive promoter that is stably transfected into a G-CSF–responsive cell line. We isolated a drug-resistant clone of the murine myeloid cell line NFS60 that contained a G-CSF–responsive reporter construct consisting of four copies of a synthetic STAT-binding element linked to a minimal promoter and the gene for luciferase (*11*). This clone, 4B6, exhibited a 20-fold increase in luciferase activity in response to G-CSF and a pattern of JAK and STAT activation similar to that seen in the parental NFS60 cells (*9*). For screening, 4B6 cells were exposed to individual synthetic organic compounds at a concentration of 10 μM, and one compound, SB 247464 (Fig. 1), was selected for further study. In the luciferase assay, SB 247464 (1 μM) had an efficacy 30% of that of G-CSF and exhibited a biphasic dose-response curve (Fig. 2). The activity of SB 247464 was evaluated on a second NFS60-based stable cell line, RSVluc, which contains stably integrated copies of a reporter plasmid that produces luciferase constitutively (*12*). The level of luciferase activity in this line is not affected by G-CSF or SB 247464 (*13*). Likewise, SB 247464 had no effect in stable cell lines containing STAT-responsive reporters that increase luciferase activity in response to either erythropoietin, interferon α, or interferon γ (*13*).

We tested whether SB 247464 caused activation of signal transduction pathways normally activated by G-CSF. Proteins from lysates of NFS60 cells treated with SB 247464 or G-CSF were precipitated with antisera to JAK1, JAK2, G-CSF receptor, STAT3, or STAT5 and detected with an antibody to phosphotyrosine. Like G-CSF, SB 247464 caused tyrosine phosphorylation of both



SB 247464

Fig. 1. Structure of SB 247464. The benzimidazole groups are arbitrarily shown in the *trans* configuration.

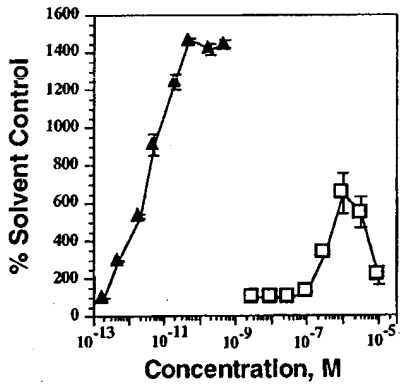


Fig. 2. Activity of G-CSF and SB 247464 in NFS60 cell luciferase assays. Dose-response curves are shown for G-CSF and SB 247464 in NFS60 cells containing a G-CSF–responsive reporter. Cytokine-independent NFS60 cells containing a stably integrated G-CSF–responsive luciferase reporter plasmid were plated in 96-well plates in Roswell Park Memorial Institute (RPMI) 1640 media containing fetal bovine serum (FBS) (0.5%), then treated with the indicated concentration of human G-CSF (▲) or SB 247464 (□) in the presence of 0.1% dimethyl sulfoxide (DMSO) for 2.5 hours. Cells were lysed, and luciferase activity was measured. All determinations were made in triplicate.

S.-S. Tian, P. Lamb, L. Kessler, J. Rosen, Department of Transcription Research, Ligand Pharmaceuticals, San Diego, CA 92121, USA. A. G. King, L. Averill, L. M. Pelus, S. B. Dillon, Department of Molecular Virology and Host Defence–US, SmithKline Beecham Pharmaceuticals, Collegeville, PA 19426, USA. S. G. Miller, Department of New Leads Discovery, Ligand Pharmaceuticals, San Diego, CA 92121, USA. J. I. Luengo and J. G. Gleason, Department of Medicinal Chemistry–US, SmithKline Beecham Pharmaceuticals, Collegeville, PA 19426, USA. R. K. Johnson, Department of Oncology, SmithKline Beecham Pharmaceuticals, King of Prussia, PA 19406, USA.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: plamb@ligand.com

JAK1 and JAK2 (Fig. 3A). SB 247464 also caused tyrosine phosphorylation of the G-CSF receptor, but not the interleukin-3 (IL-3) receptor (Fig. 3B). Both G-CSF and SB 247464 induced tyrosine phosphorylation of STAT3 and STAT5 (Fig. 3C). The two bands that became tyrosine phosphorylated in the STAT3 immunoprecipitations represent STAT3 isoforms (*14*). The time course of STAT activation in response to SB 247464 or G-CSF was very similar.

We tested whether ectopic expression of the murine G-CSF receptor was sufficient to confer sensitivity to SB 247464 on nonresponsive cells. The human megakaryocytic cell line UT7Epo (*15*) does not express the G-CSF receptor, and STATs were not activated after G-CSF or SB 247464 treatment (Fig. 4A). However, UT7Epo cells stably transfected with an expression vector containing the murine G-CSF receptor cDNA (*10, 16*) became responsive to both G-CSF and SB 247464, demonstrated by the induction of STAT-DNA complexes (Fig. 4A). Transfection of the human hepatoma cell line HepG2 (*17*) with a murine G-CSF receptor expression vector also conferred sensitivity to either G-CSF or SB 247464, as measured using a STAT-responsive luciferase reporter (Fig. 4B).

In a number of human myeloid G-CSF–responsive cell lines, SB 247464 failed to show measurable activity as judged by induction of activated STATs or G-CSF early response genes (*13*). This is in contrast to the lack of mouse-human species specificity exhibited by G-CSF itself. We exploited the species specificity of SB 247464 to determine whether it requires the extracellular or intracellular domain of the murine G-CSF receptor for activity. A chimeric murine-human G-CSF receptor was constructed by replacing the sequences encoding the extracellular domain and part of the

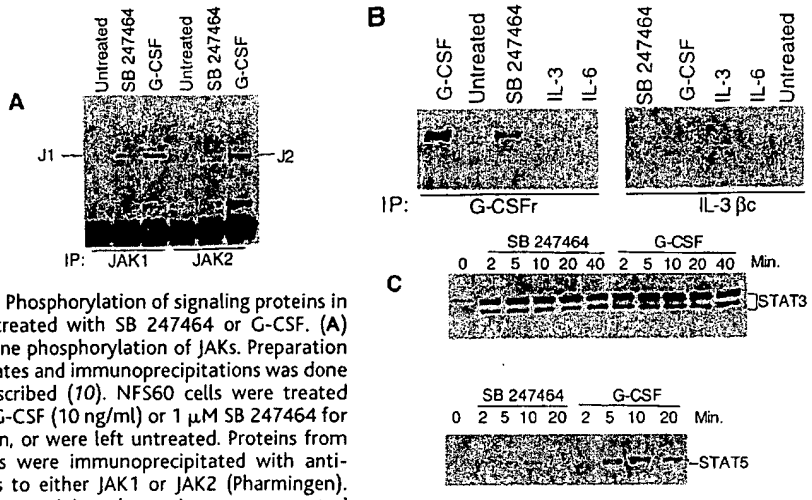


Fig. 3. Phosphorylation of signaling proteins in cells treated with SB 247464 or G-CSF. (**A**) Tyrosine phosphorylation of JAKs. Preparation of lysates and immunoprecipitations was done as described (*10*). NFS60 cells were treated with G-CSF (10 ng/ml) or 1 μM SB 247464 for 10 min, or were left untreated. Proteins from lysates were immunoprecipitated with antibodies to either JAK1 or JAK2 (Pharmingen). Immunoprecipitated proteins were separated by electrophoresis (8% gel), blotted onto a membrane, and detected with an antibody to phosphotyrosine (4G10, Upstate Biotech, Lake Placid, NY). (**B**) Tyrosine phosphorylation of the G-CSF receptor. NFS60 cells were treated for 10 min with G-CSF (10 ng/ml), IL-3 (10 ng/ml), IL-6 (10 ng/ml), or 1 μM SB 247464 in RPMI containing FBS (0.5%) and 0.1% DMSO. Lysates were made and processed as in (A), except that a polyclonal antisera to the cytoplasmic domain of the murine G-CSF receptor (*15*) or to the β-chain of the IL-3 receptor (Santa Cruz) was used for immunoprecipitations. (**C**) Tyrosine phosphorylation of STAT proteins. NFS60 cells in RPMI containing FBS (0.5%) were treated with G-CSF (10 ng/ml) or 1 μM SB 247464 for the indicated times, or were left untreated. Lysates were prepared and immunoprecipitated as described in (A) with antibodies to either STAT3 (J. Darnell, Upstate Biotechnology) or STAT5 (Santa Cruz).

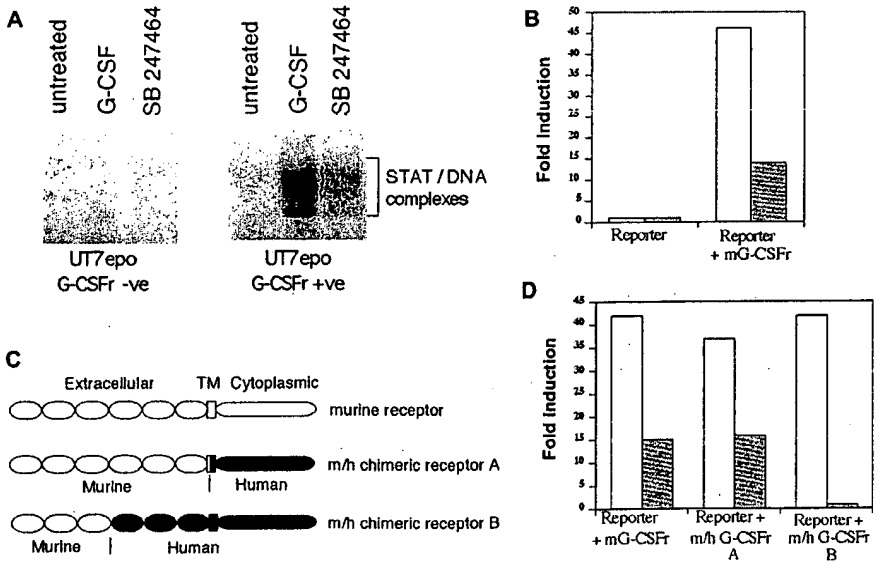


Fig. 4. The murine G-CSF receptor confers responsiveness to SB 247464. (**A**) UT7Epo cells (G-CSFr –ve) and UT7Epo cells stably transfected with the murine G-CSF receptor cDNA (G-CSFr +ve) were treated with G-CSF (10 ng/ml) or 1 μM SB 247464 for 30 min, or were left untreated. Nuclear extracts were prepared and incubated with a radiolabeled STAT-binding element, and STAT DNA complexes were separated from unbound DNA by nondenaturing gel electrophoresis. The gel was dried and exposed to x-ray film. (**B**) HepG2 cells were transfected with either 4xIRFtkluc (reporter) or with 4xIRFtkluc plus a vector directing the expression of the murine G-CSF receptor (*10, 16*) by the calcium phosphate method. Transfected cells were allowed to recover overnight, then were treated for 4 hours with either 10 ng/ml G-CSF (white bars) or 1 μM SB 247464 (gray bars). Control transfected cells were left untreated. Cells were then lysed, and luciferase levels were determined. Fold inductions were calculated by dividing the activity present in treated cells by that present in untreated cells. All transfections were performed in triplicate. (**C**) Domain structure of murine and chimeric G-CSF receptors. The location of the extracellular, cytoplasmic, and transmembrane (TM) domains are indicated. In the chimeric receptor, murine sequences are shown in white and human sequences are in black. The chimeric murine-human G-CSF receptor construct A was obtained by replacing the extracellular domain and the first 11 amino acids of the transmembrane domain of the human receptor with the Hind III (nucleotide 165) to Sca I (nucleotide 2087) fragment of the murine receptor(*16*). The chimeric receptor construct B was obtained by replacing the NH_2-terminal half of the extracellular domain of the human receptor (up to amino acid 339) with the Hind III (nucleotide 165) to Pml I (nucleotide 1199) fragment of the murine receptor. (**D**) The chimeric receptor constructs or the wild-type murine receptor construct were transfected into HepG2 cells with the 4xIRFtkluc reporter, treated, and processed as described in (B). Response to G-CSF is shown by the white bars; response to SB 247464 is shown by the gray bars.

transmembrane domain of the human receptor with the corresponding murine sequences (Fig. 4C). This chimeric construct was then transfected into HepG2 cells together with a STAT-responsive reporter. The chimeric receptor construct conferred a response to both G-CSF and SB 247464 on the HepG2 cells (Fig. 4D). This result implies that SB 247464 requires the extracellular domain of the murine G-CSF receptor for activity. A construct in which the NH_2-terminal half of the extracellular domain comprising the G-CSF binding region (*18*) is murine in origin, and the remainder of the receptor, which is human, did not confer responsiveness to SB 247464. This shows that the murine G-CSF receptor sequences required for SB 247464 activity are distinct from those required for G-CSF binding.

G-CSF normally acts on granulocytic precursor cells in the bone marrow, supporting their proliferation and differentiation. In a primary marrow colony-forming unit–granulocyte (CFU-G) assay (*19*), G-CSF and SB 247464 supported the formation of granulocytic colonies (Fig. 5). The peak efficacy of SB 247464 varied between 25 to 80% of that of G-CSF in different experiments. Colonies stimulated by SB 247464 appeared uniformly smaller than those stimulated by G-CSF, but were consistently larger than 30 cells.

Subcutaneous administration of 50 μg of G-CSF per kilogram of body weight twice a day to normal mice results in a fourfold increase in peripheral blood neutrophil counts after 4 days (Fig. 6). SB 247464 also caused a dose-dependent increase in peripheral blood neutrophils. No significant changes were noted in other blood cell populations. The efficacy of SB 247464 at 30 mg/kg is equivalent to that of 50 μg/kg of G-CSF, raising neutrophil counts approximately fourfold over baseline (Fig. 6). This fold-increase is equivalent to the increase seen when 5 to 30 μg/kg/day of G-CSF is administered to normal or neutropenic humans.

G-CSF, like other cytokines in the same family, acts by triggering dimerization or higher order oligomerization of its receptor chains (*2, 20*). The precise mechanism by which it does this is unclear, as is the mechanism by which SB 247464 is able to mimic the protein cytokine G-CSF. However, the fact that SB 247464 rapidly activates early events in the G-CSF signal transduction pathway, together with the ability of the transfected murine G-CSF receptor cDNA to confer both G-CSF and SB 247464 response to nonresponsive cells, shows that SB 247464 acts through the receptor. The twofold rotational symmetry of SB 247464 is compatible with a model in which it functions in some way as a ligand to effect dimerization of G-CSF receptor chains. This model of SB 247464 action would account for the biphasic dose response, as has been described for growth hormone (*2*). Although in longer term assays SB 247464 appears toxic at the highest concentrations used, this is not seen in short-term assays, indicating that the shape of the dose-response curve is not simply due to toxicity.

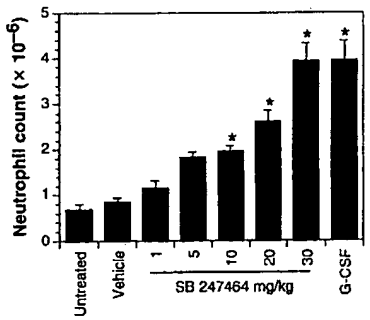


Fig. 6. Granulopoietic activity of SB 247464 in vivo. Female BDF-1 mice were given subcutaneous injections twice daily, of either G-CSF (50 μg/kg) in phosphate-buffered saline or SB 247464 dissolved in acidified H_2O (pH 4.0). Control animals received only acidified H_2O. After 4 days, blood was drawn and the number of neutrophils were counted using a Technicon hematology analyzer. Each bar represents the average of five mice; error bars show SEM. Asterisks indicate neutrophil counts that differ from those in untreated controls with a $P < 0.001$ by analysis of variance.

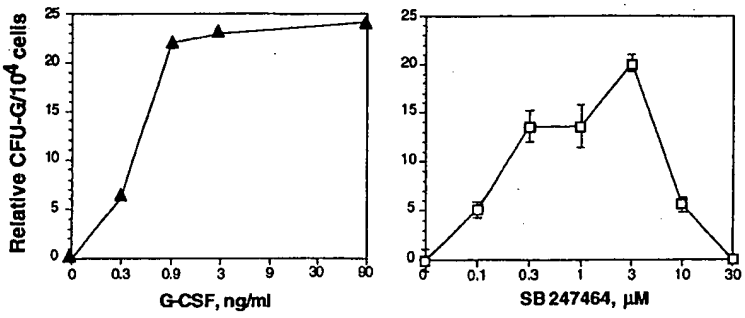


Fig. 5. Granulocytic colony formation in response to SB 247464 in vitro. Bone marrow cells (100,000 cells per 0.5 ml) obtained from female C57Bl/6 mice were incubated with either G-CSF or SB 247464 in McCoy's 5a media containing with FBS (15%) and 0.3% agar for 7 days at 37°C in a humidified incubator. Colonies of cells (>30 cells) were counted by microscopy. Data shown are the relative number of CFU-G colonies per 10^4 cells induced by each treatment. The relative number of colonies is the number of colonies in cultures of treated cells minus the number of colonies formed by control, untreated cells. Error bars indicate the standard error of the mean (SEM; $n = 6$).

The discovery of SB 247464 demonstrates that a small, nonpeptidyl molecule is capable of inducing activities normally associated with a protein hormone, both in vitro and in vivo. Our findings indicate that a small molecule can trigger the activation of a large (~120 kD) receptor protein that requires dimerization for activation, through a domain not involved in binding the natural ligand.

References and Notes

1. C.-H. Heldin, *Cell* **80**, 213 (1995).
2. J. A. Wells and A. M. d. Vos, *Annu. Rev. Biochem.* **65**, 609 (1996).
3. M. Fourcin *et al.*, *J. Biol. Chem.* **271**, 11756 (1996); T. Takahashi *et al.*, *ibid.*, p. 17555.
4. N. C. Wrighton *et al.*, *Science* **273**, 458 (1996); S. E. Cwirla *et al.*, *ibid.* **276**, 1696 (1997).
5. G. D. Demetri and J. D. Griffin, *Blood* **78**, 2791 (1991); P. Anderlini, D. Przepiorka, R. Champlin, M. Korbling, *ibid.* **88**, 2819 (1996).
6. K. Welte, J. Gabrilove, M. H. Bronchud, E. Platzer, G. Morstyn, *ibid.* **88**, 1907 (1996).
7. J. N. Ihle, B. A. Witthuhn, F. W. Quelle, K. Yamamoto, O. Silvennoinen, *Annu. Rev. Immunol.* **13**, 369 (1995).
8. S.-S. Tian, P. Lamb, H. M. Seidel, R. B. Stein, J. Rosen, *Blood* **84**, 1760 (1994); S. E. Nicholson, U. Novak, S. F. Ziegler, J. E. Layton, *ibid.* **86**, 3698 (1995); K. Shimoda *et al.*, *ibid.* **90**, 597 (1997).
9. C. Schindler and J. E. Darnell, *Annu. Rev. Biochem.* **64**, 621 (1995).
10. S.-S. Tian *et al.*, *Blood* **88**, 4435 (1996).
11. IL-3 was withdrawn from NFS60 cells. A cytokine-independent variant of NFS60 grew out, which was used for transfection experiments. The G-CSF signaling characteristics of this line are similar to the parental NFS60s. This line was transfected by electroporation with the plasmid 4xTTCC1Ctkluc, which contains four copies of the STAT-binding element 5′-GATCTGCTTCCCGGAACGT-3′ cloned into the plasmid tk-luc (*21*), and the plasmid pSV2neo, which confers neomycin resistance. Single drug-resistant clones were isolated by limiting dilution and screened for increased luciferase expression after a 2.5-hour exposure to 10 ng/ml G-CSF.
12. The plasmid RSVluc, containing the Rous sarcoma virus long terminal repeat linked to luciferase, was transfected into the cytokine-independent NFS60 line with the plasmid pSV2neo conferring neomycin resistance. Single drug-resistant clones were isolated by limiting dilution and were screened for luciferase production before and after a 2.5-hour treatment with 10 ng/ml G-CSF. A clone showing no response to G-CSF was expanded.
13. S.-S. Tian *et al.* data not shown.
14. Z. Zhong, Z. Wen, J. E. Darnell, *Science* **264**, 95 (1994).
15. N. Komatsu *et al.*, *Blood* **82**, 456 (1993).
16. R. Fukunaga, E. Ishizaka-Ikeda, Y. Seto, S. Nagata, *Cell* **61**, 341 (1990).
17. D. P. Aden, A. Fogel, S. Plotkin, I. Damjanov, B. B. Knowles, *Nature* **282**, 615 (1979).
18. R. Fukunaga, E. Ishizaka-Ikeda, C.-X. Pan, Y. Seto, S. Nagata, *EMBO J.* **10**, 2855 (1991).
19. D. Metcalf and N. A. Nicola, *J. Cell. Physiol.* **116**, 198 (1983).
20. T. Horan *et al.*, *Biochemistry* **35**, 4886 (1996); O. Hiraoka, H. Anaguchi, Y. Ota, *FEBS Lett.* **356**, 255 (1994).
21. H. M. Seidel *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 3041 (1995).

22. We thank C. Sincich and P. Sanders for technical assistance and R. Chedester for help with the manuscript.

4 February 1998; accepted 2 June 1998

EXHIBIT K

Commentary
Design and discovery of protein dimerizers

Paul A Clemons

Addresses
Harvard University, Departments of Molecular & Cellular Biology and Chemistry & Chemical Biology, 12 Oxford Street, Box 29, Cambridge, MA 02138, USA; e-mail: clemons@slsiris.harvard.edu

Current Opinion in Chemical Biology 1999, 3:112–115

http://biomednet.com/elecref/1367593100300112

 ISSN 1367-5931

Abbreviations

CID	chemical inducer of dimerization
FKBP12	FK506-binding protein 12
G-CSF	granulocyte colony-stimulating factor
PDGF	platelet-derived growth factor
TCR	T cell receptor

Introduction

Proximity effects are an essential part of cellular signal transduction [1,2]. Proteins regulated by inducible dimerization or co-localization are directly involved in critical processes such as cell proliferation, differentiation and apoptosis [3]. In addition to transcription factors and cytoplasmic signaling proteins, a large class of proximity-regulated molecules includes the cell-surface receptors for hormones, cytokines and growth or differentiation factors [4]. In many cases, the natural ligands for these receptors are rotationally symmetric homodimers (see Figure 1a). A recent report in *Science* [5] reveals the discovery of a novel small molecule, SB247464, which acts as a molecular dimerizer, activating one such receptor by mimicking the action of the polypeptide hormone granulocyte colony-stimulating factor (G-CSF). This commentary aims to illustrate the importance of this finding to the design and discovery of dimerizer reagents, both from a historical perspective and with an eye toward the future.

Bivalent antibodies

The most well known dimerizer reagents are the antibodies produced naturally by the immune system (Figure 1b). The molecular genetics of immunoglobulin production require that mature, functional antibodies be bivalent in their antigen-binding properties [6]. Not surprisingly, immunologists performed some of the earliest experiments using dimerizers. Antibodies to T cell receptor (TCR) subunits and other surface antigens have long been used by researchers to mimic antigen presentation or other stimuli normally provided by another cell. For example, experiments with combinations of bivalent antibodies and monovalent antibody fragments were used to show that aggregation of the TCR complex on the surface of T lymphocytes is necessary and sufficient to stimulate the calcium flux required for T cell activation [7].

Studies using antibodies as receptor agonists have identified proximity effects that, depending upon the cell type and circumstances, result in either induction or inhibition of cell proliferation. In a prolactin-dependent cell line, antibodies to the prolactin receptor were sufficient to induce phosphorylation of a receptor-associated kinase, JAK2 (Janus kinase 2), and to promote cell proliferation [8]. Autocrine proliferation of a squamous cell carcinoma line was inhibited by antibodies to the epidermal growth factor receptor. This inhibition is a consequence of receptor dimerization and internalization, though it does not seem to require receptor phosphorylation [9]. Human erythroid cell precursors formed erythroid colonies in culture in response to antibodies to the erythropoietin receptor [10]. In each of these cases, monovalent fragments derived from the antibodies failed to mediate similar effects.

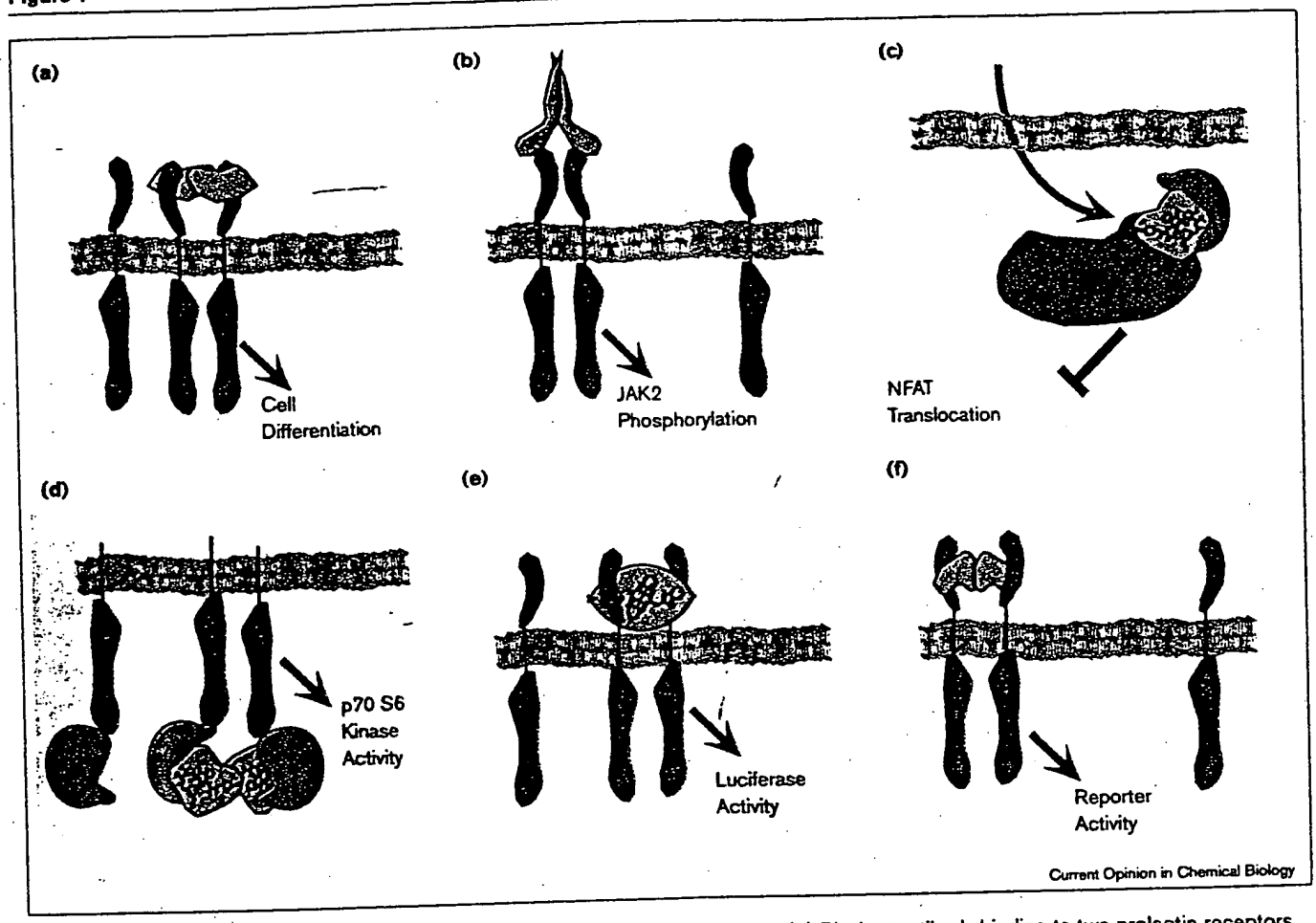
While these studies support dimerization as a mechanism for activating various growth and differentiation factor receptors, the utility of antibodies as molecular dimerizers is limited. The production of antibodies is relatively time-consuming, and the generation of highly specific antibodies that activate a particular receptor is by no means guaranteed. Furthermore, as both proteins and immunogens, antibodies are potentially less useful therapeutically than small organic molecules (the immune system is specifically geared to mount a response to protein antigens). This final limitation is similarly applicable to the use of protein fragments or peptides as receptor ligands, despite much promising work in this area [11].

Chemical inducers of dimerization

Three small-molecule immunosuppressants, FK506, cyclosporin and rapamycin, represent another class of naturally occurring dimerizers (Figure 1c). These compounds each mediate the association of a small protein, called an immunophilin, with a cellular regulatory protein. An interesting three-way relationship exists between these heterodimerizers. FK506 binds FK506-binding protein 12 (FKBP12), and this complex subsequently binds the cellular phosphatase calcineurin, inhibiting its enzymatic activity. This inhibition abolishes the ability of calcineurin to promote nuclear translocation of NFAT (nuclear factor of activated T cells), a transcription factor required for lymphocyte activation. Calcineurin phosphatase activity is also inhibited by cyclosporin, but only in complex with an independent immunophilin, cyclophilin [12]. Like FK506, rapamycin mediates its immunosuppressive effects in complex with FKBP12, but in this case, the complex binds and inhibits the protein kinase FRAP (FKBP12–rapamycin-associated protein) [13].

Since the elucidation of their mechanisms, these chemical inducers of dimerization (CIDs) have provided valuable information about the signaling pathways in which both

Figure 1



Schematic representation of molecular dimerizers. Representations of various dimerizer reagents discussed in the text are depicted. In each case, the dimerizer is shown in lime-green, the extracellular domain of the receptor in blue, and the intracellular domain in red. Specific experimental readouts have been denoted to illustrate how dimerization is detected. **(a)** G-CSF dimer binding to two G-CSF receptors. **(b)** Bivalent antibody binding to two prolactin receptors. **(c)** FKBP12–FK506 binding to calcineurin. FKBP12, dark-green; calcineurin, purple. **(d)** FK1012 binding two chimeric PDGF receptors; FKBP12, dark-green. **(e)** SB247464 binding two G-CSF receptors. **(f)** Combinatorial dimerizer binding two target receptors.

calcineurin and FRAP function. Each of these molecules has been subjected to chemical derivatization to investigate the structural basis of its binding interactions. From such research has emerged a collection of small molecules with reduced or absent immunosuppressive activity. Mutagenesis strategies have proven successful in engineering the protein components of these complexes to restore binding to certain derivatized CIDs [14,15]. These so-called 'bump-hole' systems allow signal transduction to be studied without the complications associated with the inhibitory effects of the natural immunosuppressive complex. Natural and synthetic CIDs have been used to study transcriptional initiation and reinitiation, nuclear import and export, receptor dimerization, and recruitment to the plasma membrane [3].

Symmetric dimerizers

Among the CIDs employed to probe cellular proximity effects are a variety of symmetric dimers of FK506, collectively called FK1012 [16,17]. Such molecules have been used to provide conditional control over TCR activation in a manner analogous to the use of bivalent antibodies to this receptor [18]. In addition, FK1012 has been used to study the activation of the platelet-derived growth factor (PDGF) receptor (Figure 1d) [19]. HEK293 embryonic kidney cells equipped with membrane-localized fusions of FKBP12 to the cytoplasmic domain of the PDGF receptor were shown to undergo phosphorylation of the receptor chimeras, as well as activation of both mitogen-activated protein kinase and p70 S6 kinase, in a dose-dependent response to FK1012. (These biochemical events are part of the normal signaling pathway involved in PDGF activation.) In addition, FK1012 induced mesoderm formation in *Xenopus* embryo explants harboring this chemical-conditional allele of the PDGF receptor.

In the report by Tian *et al.* [5], SB247464 is first shown to activate luciferase reporter gene activity under the control of a G-CSF-responsive promoter (Figure 1e). The amplitude of such activation is about half that for G-CSF, and

the effective concentration required for half-maximal activation is several orders of magnitude higher. The internal symmetry of SB247464 and the biphasic nature of its dose-response in the luciferase assay, however, suggest a mechanism involving receptor dimerization. Furthermore, ectopic expression of the G-CSF receptor confers SB247464 sensitivity to an otherwise insensitive cell line. Experiments using chimeric human and murine G-CSF receptors demonstrate both that SB247464 is specific for the murine receptor and that ligand binding occurs extracellularly in a domain distinct from that for G-CSF binding. Finally, SB247464, like G-CSF itself, is capable of inducing granulocytic colonies in cell culture and of stimulating a fourfold increase in peripheral blood neutrophil counts in mice.

Tian *et al.* [5] demonstrated the same fundamental principle as the study of PDGF receptor signaling using FK1012. Dimerization, even induced artificially, is sufficient to mediate the activation of some cell surface receptors. In both cases, sufficient mimicry of the natural cytokine by a CID was observed to warrant use of the dimerizer to probe further the pathway in question. The underlying significance of the G-CSF study resides in the fact that SB247464 is a serendipitous dimerizer. The investigators developed a method to assay any small molecule for agonists of the G-CSF pathway. The satisfying outcome of the work of these authors [5] is that a screen of many compounds, with no bias toward symmetric molecules, identified a rotationally symmetric small molecule whose target is a member of a dimerizer-inducible superfamily of proteins. Such a result validates previous work in deliberate dimerizer design.

Combinatorial dimerizers

One of the key advantages that a molecule such as SB247464 has over one such as FK1012 is that SB247464 does not require genetically engineered alleles of the receptor in order to produce a biological response. This requirement represents one limitation in the design of dimerizers based upon immunosuppressant–immunophilin interactions. While genetically engineered proteins allow tissue specificity to be addressed more directly, not all fusion proteins involving immunophilins will fold or function properly, nor will all active fusion proteins dimerized in this way lead to physiologically meaningful interactions. The fact that SB247464 apparently acts via receptor dimerization suggests the synthetic attainability of a wide range of receptor dimerizers. What is needed is a mechanism for the design and discovery of other such molecules, including the means to screen such molecules in functional cellular assays.

Recent efforts both in the generation of recombinant peptide libraries and in the synthesis of combinatorial libraries of natural product-like molecules suggest an avenue for the production of dimerizer libraries (Figure 1f) [11,20,21]. In such a system, a rotationally symmetric template might be attached to the solid phase, then subjected to multiple rounds of chemical transformation using panels of building blocks, each round preceded by a pool-split step to ensure the maximum diversity of products [22,23]. The resulting molecules would each be bivalent, since attachment to the solid phase would result in exposure of symmetric moieties to the same set of reagents in each round [24]. By generating a large number of molecules in this way, we improve our chances of finding specific agonists for cell surface receptors.

A complementary problem to dimerizer ligand design is the development of assay methods to screen extremely large libraries of synthetic compounds with high efficiency and throughput [25–28]. In the case of the SB247464 report [5], the assay was developed using a reporter gene driven by a promoter known to be responsive to the growth factor of interest. In principle, this provides a good way to identify activators of a particular pathway, but optimally requires the unlikely one-to-one correspondence between promoter and pathway. One hopeful note regarding assay development for receptor agonists is the idea that full receptor occupancy is unlikely to be required in order to achieve a detectable signal. In other words, many readouts (such as reporter gene activation) probably require only a small fraction of a receptor population to be dimerized due to signal amplication by the cell. Nevertheless, many receptors and their dimerization-inducible processes remain poorly characterized, suggesting the need for more general screens that do not require advance knowledge of the details of a particular pathway.

Conclusions

Beginning with the historical efforts in immunology to cross-link cell surface receptors as a means of studying lymphocyte activation, dimerizer research has grown to encompass a diverse set of tools for inducing proximity between proteins. Natural growth factors have been winnowed or emulated in the search for peptides with dimerizer properties, though the caveats associated with such immunogenic molecules remain a concern. Inspired by the natural immunosuppressant heterodimerizers, both nonimmunosuppressive derivatives, as well as synthetic homodimerizers, have been designed. While these molecules have shed light on several biological processes in cell culture, their ultimate utility in animals is limited by the need to engineer proteins for responsiveness to these ligands.

The discovery of SB247464 heralds yet another phase in dimerizer design, providing for the first time conclusive evidence that a small nonpeptidyl ligand, prepared by organic synthesis, can mimic the effects of a natural growth factor. In the future, screens for such agonists will involve probes for more general readouts of receptor activation, such as differences in post-translational modification, readily detectable morphological changes, or effects on proliferation or survival. Assay methods based on immunodetection may return us to the roots of research involving dimerizer reagents. This time, however, we are armed with

the results of many valuable studies that have laid the groundwork for preparing a new set of molecular tools.

Acknowledgements

Research by the author on small-molecule dimerizers is supported by a grant from the National Institute of General Medical Sciences (GM-52067) awarded to Stuart L. Schreiber. The author would like to thank Stuart L Schreiber and members of the Schreiber group for their invaluable contributions to these studies, and Randall T Peterson and Brent R Stockwell for their critical evaluation of the manuscript.

References

1. Seed B: Making agonists of antagonists. *Chem Biol* 1994, 1:125-129.
2. Austin DJ, Crabtree GR, Schreiber SL: Proximity versus allostery: the role of regulated protein dimerization in biology. *Chem Biol* 1994, 1:131-136.
3. Klemm JD, Schreiber SL, Crabtree GR: Dimerization as a regulatory mechanism in signal transduction. *Annu Rev Immunol* 1998, 16:569-592.
4. Heldin CH: Dimerization of cell surface receptors in signal transduction. *Cell* 1995, 80:213-223.
5. Tian SS, Lamb P, King AG, Miller SG, Kessler L, Luengo JI, Averill L, Johnson RK, Gleason JG, Pelus LM *et al.*: A small, nonpeptidyl mimic of granulocyte-colony-stimulating factor. *Science* 1998, 281:257-259.
6. Abbas A, Lichtman A, Pober J: *Cellular and Molecular Immunology*, edn 2. New York: WB Saunders Company; 1994.
7. Ratcliffe MJ, Coggeshall KM, Newell MK, Julius MH: T cell receptor aggregation, but not dimerization, induces increased cytosolic calcium concentrations and reveals a lack of stable association between CD4 and the T cell receptor. *J Immunol* 1992, 148:1643-1651.
8. Rui H, Lebrun JJ, Kirken RA, Kelly PA, Farrar WL: JAK2 activation and cell proliferation induced by antibody-mediated prolactin receptor dimerization. *Endocrinology* 1994, 135:1299-1306.
9. Fan Z, Lu Y, Wu X, Mendelsohn J: Antibody-induced epidermal growth factor receptor dimerization mediates inhibition of autocrine proliferation of A431 squamous carcinoma cells. *J Biol Chem* 1994, 269:27595-27602.
10. Elliott S, Lorenzini T, Yanagihara D, Chang D, Elliott G: Activation of the erythropoietin (EPO) receptor by bivalent anti-EPO receptor antibodies. *J Biol Chem* 1996, 271:24691-24697.
11. Dower WJ: Targeting growth factor and cytokine receptors with recombinant peptide libraries. *Curr Opin Chem Biol* 1998, 2:328-334.
12. Liu J, Farmer JD Jr, Lane WS, Friedman J, Weissman I, Schreiber SL: Calcineurin is a common target of cyclophilin–cyclosporin A and FKBP–FK506 complexes. *Cell* 1991, 66:807-815.
13. Brown EJ, Albers MW, Shin TB, Ichikawa K, Keith CT, Lane WS, Schreiber SL: A mammalian protein targeted by G1-arresting rapamycin–receptor complex. *Nature* 1994, 369:756-758.
14. Belshaw PJ, Schoepfer J, Liu K, Morrison K, Schreiber SL: Rational design of orthogonal receptor–ligand combinations. *Angew Chem Int Ed Engl* 1995, 34:2129-2132.
15. Liberles SD, Diver ST, Austin DJ, Schreiber SL: Inducible gene expression and protein translocation using nontoxic ligands identified by a mammalian three-hybrid screen. *Proc Natl Acad Sci USA* 1997, 94:7825-7830.
16. Spencer DM, Wandless TJ, Schreiber SL, Crabtree GR: Controlling signal transduction with synthetic ligands. *Science* 1993, 262:1019-1024.
17. Diver ST, Schreiber SL: Single-step syntheses of cell permeable protein dimerizers that activate signal transduction and gene expression. *J Am Chem Soc* 1997, 119:5106-5109.
18. Pruschy MN, Spencer DM, Kapoor TM, Miyake H, Crabtree GR, Schreiber SL: Mechanistic studies of a signaling pathway activated by the organic dimerizer FK1012. *Chem Biol* 1994, 1:163-172.
19. Yang J, Symes K, Mercola M, Schreiber SL: Small-molecule control of insulin and PDGF receptor signaling and the role of membrane attachment. *Curr Biol* 1998, 8:11-18.
20. Tan DS, Foley MA, Shair MD, Schreiber SL: Stereoselective synthesis of over two million compounds having structural features reminiscent of natural products and compatible with miniaturized cell-based assays. *J Am Chem Soc* 1998, 120:8565-8566.
21. Boger DL, Goldberg J, Jiang W, Chai W, Ducray P, Lee JK, Ozer RS, Andersson CM: Higher order iminodiacetic acid libraries for probing protein-protein interactions. *Bioorg Med Chem* 1998, 6:1347-1378.
22. Furka A, Sebestyen F, Asgedom M, Dibo G: General method for rapid synthesis of multicomponent peptide mixtures. *Int J Pept Protein Res* 1991, 37:487-493.
23. Lam KS, Salmon SE, Hersh EM, Hruby VJ, Kazmierski WM, Knapp RJ: A new type of synthetic peptide library for identifying ligand-binding activity. *Nature* 1991, 354:82-84.
24. Schreiber SL, Goulet MT, Schulte G: Two-directional chain synthesis of *syn*-skipped polyol chains from meso precursors. *J Am Chem Soc* 1987, 109:4718-4720.
25. Borchardt A, Liberles SD, Biggar SR, Crabtree GR, Schreiber SL: Small molecule-dependent genetic selection in stochastic nanodroplets as a means of detecting protein–ligand interactions on a large scale. *Chem Biol* 1997, 4:961-968.
26. You AJ, Jackman RJ, Whitesides GM, Schreiber SL: A miniaturized arrayed assay format for detecting small molecule–protein interactions in cells. *Chem Biol* 1997, 4:969-975.
27. Huang J, Schreiber SL: A yeast genetic system for selecting small molecule inhibitors of protein–protein interactions in nanodroplets. *Proc Natl Acad Sci USA* 1997, 94:13396-13401.
28. Stockwell BR, Haggarty SJ, Schreiber SL: High throughput post-transcriptional and biosynthetic profiling using small molecules in miniaturized mammalian cell based assays. *Chem Biol* 1999, 6:71-83.